407/B.9.7

THE AMERICAN JOURNAL OF PSYCHOLOGY

EDITED BY

KARL M. DALLENBACH
UNIVERSITY OF TEXAS

AND

M. E. BITTERMAN
BRYN MAWR COLLEGE

E. B. NEWMAN
HARVARD UNIVERSITY

LEO POSTMAN
UNIVERSITY OF CALIFORNIA, BERKELEY

WITH THE COÖPERATION OF

E. G. BORING, Harvard University; W. K. ESTES, Stanford University; J. P. GUILFORD, University of Southern California; HARRY HELSON, Kansas State University; E. R. HILGARD, Stanford University; FRANCIS W. IRWIN, University of Pennsylvania; F. NOWELL JONES, University of California, Los Angeles; G. L. KREEZER, Washington University; D. G. MARQUIS, Massachusetts Institute of Technology; ARTHUR W. MELTON, University of Michigan; W. C. H. PRENTICE, Wheaton College; T. A. RYAN, Cornell University.

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY

VOL. LXXVI
MEZES HALL, THE UNIVERSITY OF TEXAS
AUSTIN, TEXAS

Copyright 1963 by Karl M. Dallenbach

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated 8.9.70
Acs. No. J403

TABLE OF CONTENTS

Articles and Notes

Alfert, E., 'Spread of Effect' without Guessing Sequences 638

Alluisi, E. A., The Fifty-Fifth Annual Meeting of the Southern Society for Philosophy and Psychology 512

Bakan, P., and Thompson, R., Attention and Kinesthetic After-Effects 39

Battig, W. F., and Brackett, H. R., Method of Pretraining and Knowledge of Results in Paired-Associate Learning under Conditions of Repetition and Non-Repetition 66

Baxter, R., and Kimmel, H. D., Conditioning and Extinction in the Planarian 665

Beare, A. C., Color-Name as a Function of Wave-Length 248

Berg, H. W., Hansen, B., and Pangborn, R. M., The Influence of Color on Discrimination of Sweetness in Dry Table Wine 492

Berry, R. N., On Lotsof's Proposed Relation between Certainty and Expectancy 333

Bevan, W., Helson, H., and Maier, R. A., The Influence of Context upon the Estimation of Number 464

Biersdorf, W. R., Kozil, D. J., and Ohwaki, S., The Effect of Instructions and Oculomotor Adjustments on Apparent Size 1

Bilodeau, E. A., and Johnson, J. W., Task, Task-Intensity, and Dollars Demanded for Performance 293

Bitterman, M. E., Bullock, D. H., and Roberts, W. A., Resistance to Extinction in the Pigeon after Partially Reinforced Instrumental Training under Discrete-Trials Conditions 353

Bitterman, M. E., Eskin, R. M., and Gonzalez, R. C., Further Experiments on Partial Reinforcement in the Fish 366

Boardman, W. K., and Gunter, R. C., Jr., A Flexible Four-Field Tachistoscope 503

Bolen, C. H., Nicora, B. D., and Zimmer, H. A High-Speed Constant-Illumination Tachistoscope 691

Boring, E. G., The 1963 Meeting of the American Philosophical Society .. 336

Boring, E. G., The 1963 Meeting of the National Academy of Sciences 336

Bourne, L. E., Jr., Factors Affecting Strategies Used in Problems of Concept-Formation 229

Bowman, R. E., and Thurlow, W. R., Determinants of the Effect of Position in Serial Learning 436

Brackett, H. R., and Battig, W. F., Method of Pretraining and Knowledge of Results in Paired-Associate Learning under Conditions of Repetition and Non-Repetition 66

Brooks, V., and Hochberg, J., Graphic Symbols: Things or Holes? 326

Brown, C. C., and Whitman, J. R., Apparatus for Measuring Restricted Ranges of Linear Movement 138

Bryan, J. D., Milauckas, S. W., and Young, R. K., Serial Learning as a Function of Prior Paired-Associate Training 82

Bullock, D. H., Bitterman, M. E., and Roberts, W. A., Resistance to Extinction in the Pigeon after Partially Reinforced Instrumental Training under Discrete-Trials Conditions 353

Burright, R. G., Tromater, L. J., and Young, P. T., Preferences of the White Rat for Solutions of Sucrose Hydrochloride 205

Cardno, J. A., Experimental and Physiological Psychology: Historical Relations 509

CARP, F. M., and EWART, A. G., Recognition of Tactual Form by Sighted and Blind Subjects 488
CAUTELA, J. R., and MIKAELIAN, H., A Machine Capable of Being Conditioned 128
CHAMPION, R. A., and FUREDY, J. J., Cognitive and S-R Interpretations of Incentive-Motivational Phenomena 616
CHAMPION, R. A., and HEATHCOTE, M. J., A Test of the Relative Valence Theory of Instrumental Reward Learning 679
CHENZOFF, A. P., GREGG, L. W., and LAUGHERY, K. R., The Effect of Rate of Presentation, Substitution and Mode of Response in Paired-Associate Learning 110
CLARK, B., and GRAYBIEL, A., Contributing Factors in the Perception of the Oculogravic Illusion 18
COLEMAN, E. B., Approximations to English: Some Comments on the Method 239
COPPOLA, L. A., and SCHUCK, J. R., Contact Occluders for Restricting Vision in the Rat 135
COSTELLO, C. G., The Effects of Stimulant and Depressant Drugs on Physical Persistence 698
COSTELLO, C. G., and STEPHEN, J., The Tape-Recorder as an Instrument for Measuring Time 324
COSTELLO, C. G., HERRERA, B., and HOLLAND, H. C. The Role of the Interpolated Stimulus in Producing Kinesthetic Figural After-Effects 670
CROSS, H. A., and ROSENBLUM, L. A., Performance of Neonatal Monkeys in the Visual Cliff-Situation 318
CROVITZ, H. F., and LIPSCOMB, D. B., Dominance of the Temporal Visual Fields at a Short Duration of Stimulus 631
DALLENBACH, K. M., A Pain-Scale: A Report from a Victim of Tic Doloureux 152
DALLENBACH, K. M., Erratum 158
DALLENBACH, K. M., Tables vs. Graphs as Means of Presenting Experimental Results 700
DAVENPORT, J. W., The Multipurpose Psychological Laboratory at Amherst College 694
DAVIDSON, L., MOORE, M. E., and WALLACH, H., Modification of Stereoscopic Depth-Perception 191
DAWSON, W. W., LINDNER, R. S., and PERRY, N., The Electrical Quantification of the Consumption of Liquids by Rats 687
DENNER, B., MCFARLAND, J. H., WAPNER, S., and WERNER, H., Rhythmic Activity and the Perception of Time 287
DICK, E. D., Formation and Generalization of Concepts as a Function of Similarity and Amount of Training 607
DI LOLLO, V., and MARSHALL, A. J., Hering's Illusion with Impoverishment of the Stimulus in Scotopic and Photopic Vision 644
DREGER, R. M., The Ninth Annual Meeting of the Southeastern Psychological Association 703
DYAL, J. A., and HOLLAND, T. A., Resistance to Extinction as a Function of Number of Reinforcements 332
DYAL, J. A., and STABLER, J. R., Discriminative Reaction-Time as a Joint Function of Manifest Anxiety and Intelligence 484
ELITHORN, A., JONES, D., and KERR, M., A Binary Perceptual Maze 506
ELLIOTT, L. L., and TANNENBAUM, P. H., Factor-Structure of Semantic Differential Responses to Visual Forms and Prediction of Factor-Scores from Structural Characters of the Stimulus-Shapes 589
ENGEN, T., Cross-Adaptation to the Aliphatic Alcohols 96
EPSTEIN, W., The Influence of Assumed Size on Apparent Distance 257
ERICKSON, C. C., INGRAM, R. D., and YOUNG, R. K., Paired-Associate Learning as a Function of Rate of Presentation and Prior Serial Learning 458
ERVIN, S. M., and LANDAR, H., Navaho Word-Association 49
ESKILDSEN, P. R., An Apparatus for Studies of Brightness which Controls

Variations of Intensity with Time 321
ESKIN, R. M., BITTERMAN, M. E., and GONZALEZ, R. C., Further Experiments on Partial Reinforcement in the Fish 366
EWART, A. G., and CARP, F. M., Recognition of Tactual Form by Sighted and Blind Subjects 488
FILLENBAUM, S., Contextual Effects in Judgment as a Function of Restrictions in Response-Language 103
FRENCH, G. M., and MALMSTROM, E. J., Scale-Symmetry and the Semantic Differential 446
FUREDY, J. J., and CHAMPION, R. A., Cognitive and S-R Interpretations of Incentive-Motivational Phenomena 616
FURTH, H. G., and MENDEZ, R. A., The Influence of Language and Age on Gestalt Laws of Perception 74
GIBSON, J. J., and PICK, A. D., Perception of Another Person's Looking Behavior 386
GOGEL, W. C., HARKER, G. S., and WIST, E. R., A Test of the Invariance of the Ratio of Perceived Size to Perceived Distance 537
GONZALEZ, R. C., BITTERMAN, M. E., and ESKIN, R. M. Further Experiments on Partial Reinforcement in the Fish 366
GRAYBIEL, A., and CLARK, B., Contributing Factors in the Perception of the Oculogravic Illusion 18
GREGG, L. W., CHENZOFF, A. P., and LAUGHERY, K. R., The Effect of Rate of Presentation, Substitution and Mode of Response in Paired-Associate Learning 110
GUNTER, R. C., JR., and BOARDMAN, W. K., A Flexible Four-Field Tachistoscope 503
HANSEN, B., and PANGBORN, R. M., The Influence of Color on Discrimination of Sweetness and Sourness in Pear-Nectar 315
HANSEN, B., BERG, H. W., and PANGBORN, R. M., The influence of Color on Discrimination of Sweetness in Dry Table Wine 492
HARKER, G. S., GOGEL, W. C., and WIST, E. R., A Test of the Invariance of the Ratio of Perceived Size to Perceived Distance 537
HARTMAN, A. M., Effects of Retinal Figure-Ground Ratios on Target-Size and Apparent Size of After-Images 120
HEATHCOTE, M. J., and CHAMPION, R. A., A Test of the Relative Valence Theory of Instrumental Reward Learning 679
HEINEMANN, E. G., and RUDOLPH, R. L., The Effect of Discriminative Training on the Gradient of Stimulus-Generalization 653
HELD, R., Localized Normalization of Tilted Lines 146
HELSON, H., BEVAN, W., and MAIER, R. A., The Influence of Context upon the Estimation of Number 464
HERRERA, B., COSTELLO, C. G., and HOLLAND, H. C., The Role of Interpolated Stimulus in Producing Kinesthetic Figural After-Effects 670
HINCKLEY, E. D., and RETHLINGSHAFER, D., Influence of Judges' Characteristics upon the Adaptation-Level 116
HOCHBERG, J., and BROOKS, V., Graphic Symbols: Things or Holes? 326
HOLLAND, H. C., COSTELLO, C. G., and HERRERA, B., The Role of Interpolated Stimulus in Producing Kinesthetic Figural After-Effects 670
HOLLAND, T. A., and DYAL, J. A., Resistance to Extinction as a Function of Number of Reinforcements 332
HOWARD, I. P., and TEMPLETON, W. P., A Critical Note on the Use of the Human Centrifuge 150
HOWARD, T. C., The Relation between Psychological and Mathematical Probability 335
INGRAM, R. D., ERICKSON, C. C., and YOUNG, R. K., Paired-Associate Learning as a Function of Rate of Presentation and Prior Serial Learning 458
IVERSON, M. A., Thirty-Fourth Annual Meeting of the Eastern Psychological Association 514

JOHNSON, J. W., and BILODEAU, E. A., Task, Task-Intensity, and Dollars Demanded for Performance 293
JONES, D., ELITHORN, A., and KERR, M., A Binary Perceptual Maze 506
KAHNEMAN, D., The Semantic Differential and the Structure of Inferences among Attributes 554
KANUNGO, R., and LAMBERT, W. E., Semantic Satiation and Meaningfulness 421
KARSH, E. B., and WALLACH, H. Why the Modification of Stereoscopic Depth-Perception Is So Rapid 413
KARSH, E. B., and WALLACH, H., The Modification of Stereoscopic Depth-Perception and the Kinetic Depth-Effect 429
KENNEDY, J. L., Fifty-Ninth Annual Meeting of the Society of Experimental Psychologists 704
KENYON, J., and MÜLLER, H. F., A Visual Rating Scale for Prints of Finger-Sweat 140
KEPPEL, G., and UNDERWOOD, B. J., Bidirectional Paired-Associate Learning 470
KERR, M., ELITHORN, A., and JONES, D., A Binary Perceptual Maze 506
KIMMEL, H. D., and BAXTER, R., Conditioning and Extinction in the Planarian 665
KOZIL, D. J., BIERSDORF, W. R., and OHWAKI, S., The Effect of Instructions and Oculomotor Adjustments on Apparent Size 1
KRAUSKOPF, J., A Color-Mixer with Monochromatic Primaries 496
KRAVITZ, J. H., LINDAUER, J., and WALLACH, H., A Passive Condition for Rapid Adaptation to Displaced Visual Direction 554
LAMBERT, W. E., and KANUNGO, R., Semantic Satiation and Meaningfulness 421
LANDAR, H., and ERVIN, S. M., Navaho Word-Association 49
LAUGHERY, K. R., CHENZOFF, A. P., and GREGG, L. W., The Effect of Rate of Presentation, Substitution and Mode of Response in Paired-Associate Learning 110
LERNER, R. A., and WINNICK, W. A., Intentional and Incidental Learning under Distraction 683
LINDAUER, J., KRAVITZ, J. H., and WALLACH, H., A Passive Condition for Adaptation to Displaced Visual Direction 554
LINDNER, R. S., DAWSON, W. W., and PERRY, N., The Electrical Quantification of the Consumption of Liquids by Rats 687
LIPSCOMB, D. B., and CROVITZ, H. F., Dominance of the Temporal Visual Field at a Short Duration of Stimulus 631
LOGAN, F. A., and WAGNER, A. R., Choice-Behavior and Distribution of Reinforcement 480
LURIA, S. M., The Effect of Body-Position on Meridional Variation in Scotopic Acuity 598
MCFARLAND, J. H., DENNER, B., WAPNER, S., and WERNER, H., Rhythmic Activity and the Perception of Time 287
MAIER, R. A., BEVAN, W., and HELSON, H., The Influence of Context upon the Estimation of Number 464
MALMSTROM, E. J., and FRENCH, G. M., Scale-Symmetry and the Semantic Differential 446
MARSHALL, A. J., and DI LOLLO, V., Hering's Illusion with Impoverishment of the Stimulus in Scotopic and Photopic Vision 644
MARSHALL, M. A., and RUNQUIST, W. N., Transfer, Synonymity, and Anticipatory Interval in Paired-Associate Verbal Learning 281
MARTIN, J. G., and NELSON, J. T., Recognition-Tests, Repetition, and All-or-None Learning 675
MENDEZ, R. A., and FURTH, H. G., The Influence of Language and Age on Gestalt Laws of Perception 74
MIKAELIAN, H., and CAUTELA, J. R., A Machine Capable of Being Conditioned 128
MILAUCKAS, E. W., BRYAN, D. J., and YOUNG, R. K., Serial Learning as a Function of Prior Paired-Associate Training 82
MILLER, G. A., Ronald Aylmer Fisher: 1890-1962 157

Miller, W. C., Thirty-Third Annual Meeting of the Rocky Mountain Psychological Association 515
Moore, M. E., Davidson, L., and Wallach, H., Modification of Stereoscopic Depth-Perception 191
Müller, H. F., and Kenyon, J., A Visual Rating-Scale for Prints of Finger-Sweat 140
Murray, C. S., and Pierrel, R., Some Relationships between Comparative Judgment, Confidence, and Decision-Time in Weight-Lifting 28
Nealey, S. M., and Riley, B. A., Loss and Recovery of Discrimination of Visual Depth in Dark-Reared Rats 329
Neisser, U., Decision-Time without Reaction-Time: Experiments in Visual Scanning 376
Nelson, J. T., and Martin, J. G., Recognition-Tests, Repetition, and All-or-None Learning 675
Newman, E. B., Seventieth Annual Meeting of the American Psychological Association 155
Nicora, B. D., Bolen, C. H., and Zimmer, H. A High-Speed-Constant-Illumiation Tachistoscope 691
Noble, C. E., and Parker, G. V. C., Experimentally Produced Meaningfulness in Paired-Associate Learning 579
Norris, C. M., and Wallach, H., Accommodation as a Distance-Cue 659
Ohwaki, S., Biersdorf, W. R., and Kozil, D. J., The Effect of Instructions and Oculomotor Adjustments on Apparent Size 1
Over, R., Size- and Distance-Estimates of a Single Stimulus under Different Viewing Conditions 452
Pangborn, R. M., and Hansen, B., The Influence of Color on Discrimination of Sweetness and Sourness in Pear-Nectar 315
Pangborn, R. M., Berg, H. W., and Hansen, B., The Influence of Color on Discrimination of Sweetness in Dry Table Wine 492
Parker, G. V. C., and Noble, C. E., Experimentally produced meaningfulness in paired-associate learning 579
Paul, C., Saltz, E., and Whitman, R. N., Performance in the Runway as a Function of Stimulus-Differentiation 124
Perloe, S. J., The Relation between Category-Rating and Magnitude-Estimation Judgments of Occupational Prestige 395
Perry, N., Dawson, W. W., and Lindner, R. S., The Electrical Quantification of the Consumption of Liquid by Rats 687
Pick, A. D., and Gibson, J. J., Perception of Another Person's Looking Behavior 386
Pierrel, R., and Murray, C. S., Some Relationships between Comparative Judgment, Confidence, and Decision-Time in Weight-Lifting 28
Rethlingshafer, D., and Hinckley, E. D., Influence of Judges' Characteristics upon the Adaptation-Level 116
Riley, D. A., and Nealey, S. M., Loss and Recovery of Discrimination of Visual Depth in Dark-Reared Rats 329
Roberts, W. A., Bitterman, M. E., and Bullock, D. H., Resistance to Extinction in the Pigeon after Partially Reinforced Instrumental Training under Discrete-Trials Conditions 353
Rollin, A. R., Shepp, B. E., and Thaller, K. E., A Modified Wisconsin General Test Apparatus for the Rat 500
Rosenblum, L. A., and Cross, H. A., Performance of Neonatal Monkeys in the Visual Cliff-Situation 318
Rudolph, R. L., and Heinemann, E. G., The Effect of Discriminative Training on the Gradient of Stimulus-Generalization 653
Runquist, W. N., and Marshall, M. A., Transfer, Synonymity, and Anticipatory Interval in Paired-Associate Verbal Learning 281
Saltz, E., Paul, C., and Whitman, R. N., Performance in the Runway as a Function of Stimulus-Differentiation 124

SCHUCK, J. R., and COPPOLA, L. A., Contact Occluders for Restricting Vision in the Rat 135
SCHWARTZ, M., Transfer from Failed Pairs as a Test of One-Trial vs. Incremental Learning 266
SHEPARD, R. N., Production of Constrained Associates and the Informational Uncertainty of the Constraint 218
SHEPP, B. E., ROLLIN, A. R., and THALLER, K. E., A Modified Wisconsin General Test Apparatus for the Rat 500
SLAMECKA, N. J., Choice Reaction-Time as a Function of Meaningful Similarity 274
STABLER, J. R., and DYAL, J. A., Discriminative Reaction-Time as a Joint Function of Manifest Anxiety and Intelligence 484
STEPHEN, J., and COSTELLO, C. G., The Tape-Recorder as an Instrument for Measuring Time 324
STEVENSON, H. W., and WEIR, M. W., The Role of Age and Verbalization in Probability-Learning 299
STRAUGHAN, J. H., The Role of Repetition in Associative Learning 148
SUTHERLAND, N. S., The Shape-Discrimination of Stationary Shapes by Octopuses 177
TANNENBAUM, P. H., and ELLIOTT, L. L., Factor-Structure of Semantic Differential Responses to Visual Forms and Prediction of Factor-Scores from Structural Characters of the Stimulus-Shapes 589
TEMPLETON, W. P., and HOWARD, I. P., A Critical Note on the Use of the Human Centrifuge 150
TERWILLIGER, R. F., Evidence for a Relationship between Figural After-Effects and After-Images 306
THALLER, K. E., ROLLIN, A. R., and SHEPP, B. E., A Modified Wisconsin General Test Apparatus for the Rat 500
THOMPSON, R., and BAKAN, P., Attention and Kinesthetic After-Effects 39
THURLOW, W. R., and BOWMAN, R. E., Determinants of the Effects of Position in Serial Learning 436
TROMATER, L. J., BURRIGHT, R. G., and YOUNG, P. T., Preferences of the White Rat for Solutions of Sucrose Hydrochloride 205
TULVING, E., Familiarity of Letter-Sequences and Tachistoscopic Identification 143
UNDERWOOD, B. J., and KEPPEL, G., Bidirectional Paired-Associate Learning 470
WAGNER, A. R., and LOGAN, F. A., Choice-Behavior and Distribution of Reinforcement 480
WALLACH, H., and KARSH, E. B., Why the Modification of Stereoscopic Depth-Perception Is So Rapid 413
WALLACH, H., and KARSH, E. B., The Modification of Stereoscopic Depth-Perception and the Kinetic Depth-Effect 429
WALLACH, H., and NORRIS, C. M., Accommodation as a Distance-Cue 659
WALLACH, H., and ZUCKERMAN, C., The Constancy of Stereoscopic Depth 403
WALLACH, H., DAVIDSON, L., and MOORE, M. E., Modification of Stereoscopic Depth-Perception 191
WALLACH, H., KRAVITZ, J. H., and LINDAUER, J., A Passive Condition for Rapid Adaptation to Displaced Visual Direction 568
WAPNER, S., DENNER, B., MCFARLAND, J. H., and WERNER, H., Rhythmic Activity and the Perception of Time 287
WEIDENFELLER, E. W., and ZIMNY, G. H., Effects of Music upon *GSR* and Heart-Rate 311
WEINSTEIN, S., The Relationship of Laterality and Cutaneous Area to Breast-Sensitivity in Sinistrals and Dextrals 475
WEIR, M. W., and STEVENSON, H. W., The Role of Age and Verbalization in Probability-Learning 299
WEISS, J. H., Further Study of the Relation between the Sound of a Word and Its Meaning 624
WERNER, H., DENNER, B., MCFARLAND, J. H., and WAPNER, S., Rhythmic

Activity and the Perception of Time 287
WHITMAN, J. R., and BROWN, C. C., Apparatus for Measuring Restricted Ranges of Linear Movement 138
WHITMAN, R. N., PAUL, C., and SALTZ, E., Performance in the Runway as a Function of Stimulus-Differentiation 124
WILSON, K. V., Multidimensional Analyses of Confusions of English Consonants 89
WINNICK, W. A., and LERNER, R. A., Intentional and Incidental Learning under Distraction 683
WIST, E. R., GOGEL, W. C., and HARKER, G. S., A Test of the Invariance of the Ratio of Perceived Size to Perceived Distance 537
WONG, R., Effect of Modulus on Estimate of Magnitude of Linear Extent ... 511
YOUNG, P. T., BURRIGHT, R. G., and TROMATER, L. J., Preferences of the White Rat for Solutions of Sucrose and Hydrochloride 205
YOUNG, R. K., BRYAN, J. D., and MILAUCKAS, E. W., Serial Learning as a Function of Prior Paired-Associate Training 82
YOUNG, R. K., ERICKSON, C. C., and INGRAM, R. D., Paired-Associate Learning as a Function of Rate of Presentation and Prior Serial Learning 458
ZIMMER, H., BOLEN, C. L., and NICORA, B. D., A High-Speed Constant-Illumination Tachistoscope 691
ZIMNY, G. H., Effects of Direction upon Speed of Mirror-Drawing 58
ZIMNY, G. H., and WEIDENFELLER, E. W., Effects of Music upon *GSR* and Heart-Rate 311
ZHCKERMAN, C., and WALLACH, H., The Constancy of Stereoscopic Depth .. 404

BOOK REVIEWS

(The reviewer's name appears in parentheses after the title of the work.)

ARNOLD, M. B., Emotion and Personality: Vol. I, Psychological Aspects; Vol. II, Neurological and Physiological Aspects (P. T. Young) 516
BAER, D. M., and BIJOU, S. W., Child Development: Vol. I, A Systematic and Empirical Theory (A. E. Siegel) 173
BIJOU, S. W., and BAER, D. M., Child Development: Vol. I, A Systematic and Empirical Theory (A. E. Siegel) 173
BJERSTEDT, A., Glimpses from the World of the School Child: Self and Other Explored in Dyadic Communication (I. D. Steiner) 168
BLUM, G. S., A Model of the Mind (P. London) 526
BORING, E. G., Psychologist at Large (T. A. Ryan) 530
BRAZIER, M. A. B. (ed.), Brain and Behavior: Vol. I, Proceedings of the First Conference (J. Orbach) 524
BUROS, O. K., Tests in Print (J. C. Nunnally) 714
BUSS, A. H., The Psychology of Aggression (A. E. Siegel) 713
CRAM, D., Explaining "Teaching Machines" and Programming (D. J. Klaus) 712
CUMMING, E., and HENRY, W. E., Growing Old: The Process of Disengagement (N. W. Shock) 341
DUIJKER, H. C. J., and FRIJDA, N. H., National Character and National Stereotypes (D. V. McGranahan) 172
DYK, R. B., *et al.,* Psychological Differentiation: Studies of Development (R. W. Gardner) 709
FATERSON, H. F., *et al.,* Psychological Differentiation: Studies of Development (R. W. Gardner) 709
FISCKE, D. W., and MADDI, S. R. (eds.), Functions of Varied Experience (D. E. Berlyne) 339
FRIJDA, N. H., and DUIJKER, H. C. J., National Character and National Stereotypes (D. V. McGranahan) 172

Goodenough, D. R., *et al.*, Psychological Differentiation: Studies of Development (R. W. Gardner) 709
Harvey, O. J., *et al.*, Conceptual Systems and Personality Organization (M. Henle) 163
Haynes, R., The Hidden Springs (C. E. M. Hansel) 170
Henry, W. E., and Cumming, E., Growing Old: The Process of Disengagement (N. W. Shock) 341
Heywood, R., Beyond the Reach of Sense: An Inquiry into Extra-Sensory Perception (C. E. M. Hansel) 170
Homans, G. C., Social Behavior: Its Elementary Forms (E. P. Hollander) .. 347
Hountras, P. T. (ed.), Mental Hygiene: A Text of Readings (R. T. Lewis) 344
Hunt, D. E., *et al.*, Conceptual Systems and Personality Organization (M. Henle) 163
Jacobs, H. H., *et al.*, Behavioral Approaches to Accident Research (T. W. Forbes) 528
Jensen, A. R., and Symonds, P. M., From Adolescent to Adult (D. B. Harris) 166
Kagan, J., and Lesser, G. S., (eds.), Contemporary Issues in Thematic Apperception Methods (W. H. Holtzman) 345
Karp, S. A., *et al.*, Psychological Differentiation: Studies of Development (R. W. Gardner) 709
Kilpatrick, F. P., (ed.), Explorations in Transactional Psychology (H. L. Gruber) 159
Klausmeier, H. J., Learning and Human Abilities: Educational Psychology (D. P. Ausubel) 165
Kluckhohn, F. R., and Strodtbeck, F. L., Variations in Value Orientation (S. M. Ervin) 342
Kuppuswamy, B., An Introduction to Social Psychology (S. Berger) 346
Lesser, G. S., and Kagan, J., (eds.), Contemporary Issues in Thematic Apperception Methods (W. H. Holtzman) 345
Levine, E. S., The Psychology of Deafness: Techniques of Appraisal for Rehabilitation (C. W. Fuller) 521
Lyle, J., *et al.*, Television in the Lives of Our Children (W. C. Olson) 711
Maddi, S. R., and Fiske, D. W. (eds.), Functions of Varied Experience (D. E. Berlyne) 339
Madison, P., Freud's Concept of Repression and Defense: Its Theoretical and Observational Language (T. A. Ryan) 337
Massarik, F., *et al.*, Leadership and Organization: A Behavioral Science Approach (L. Meltzer) 347
National Physics Laboratory, Visual Problems of Color: A Symposium (F. L. Dimmick, J. A. S. Kinney, and S. M. Luria) 519
Parker, E. B., *et al.*, Television in the Lives of Our Children (W. C. Olson) 711
Robertiello, R. C., A Handbook of Emotional Illness and Treatment (A. Ellis) 169
Robinson, F. P., Effective Study (W. Pauk) 171
Rosenblatt, F., Principles of Neurodynamics: Perception and the Theory of Brain Mechanisms (B. W. White) 705
Schramm, W., *et al.*, Television in the Lives of Our Children (W. C. Olson) 711
Schroder, H. M., *et al.*, Conceptual Systems and Personality Organization (M. Henle) 163
Scott, W. A., and Wertheimer, M., Introduction to Psychological Research (R. D. Walk) 707
Strodtbeck, F. L., and Kluckhohn, F. R., Variations in Value Orientation (S. M. Ervin) 342
Symonds, P. M., and Jensen, A. R., From Adolescent to Adult (D. B. Harris) 166
Tannenbaum, R., *et al.*, Leadership and Organization: A Behavioral Science Approach (L. Meltzer) 347
Von Békésy, G., Experiments in Hearing (W. L. Gulick) 531

Von Haller-Gilmer, B., *et al.,* Industrial Psychology (E. C. Webster) 343
Vygotsky, L. S., Thought and Language (R. B. MacLeod) 532
Wallraff, C. A., Philosophical Theory and Psychological Fact (T. A. Ryan) 161
Wertheimer, M., and Scott, W. A., Introduction to Psychological Research (R. D. Walk) 707
Weschler, I. R., *et al.,* Leadership Organization: A Behavioral Science Approach (L. Meltzer) 347
Witkin, H. A., *et al.,* Psychological Differentiation: Studies of Development (R. W. Gardner) 709

THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. STANLEY HALL

Vol. LXXVI MARCH, 1963 No. 1

THE EFFECT OF INSTRUCTIONS AND OCULOMOTOR ADJUSTMENTS ON APPARENT SIZE

By W. R. BIERSDORF, Walter Reed Army Institute of Research,
SONOKO OHWAKI, Clark University, and D. J. KOZIL,
Walter Reed Army Institute of Research

A stimulus-object of constant angular size is perceived as of different absolute size depending upon the distance at which the eyes fixate. Wheatstone, the inventor of the stereoscope, reported that an object is perceived as of smaller size when the eyes are adjusted for near vision.[1] The convergence of the eyes for near vision was held responsible for the decrease in apparent size.

Since that time, several experiments have been performed confirming the changes in apparent size with binocular viewing in stereoscopic devices and attributing the size changes to convergence and accommodation.[2] A similar effect takes place with monocular viewing. Hering demonstrated a reduction in apparent size upon changing the fixation-distance from far to near without the use of a stereoscope.[3]

Several experiments have been performed involving the perception of size through an artificial pupil which eliminates the need for the eye to change its accommodation in order to see clearly at various distances. Maier found also that an artificial pupil

---

* Received for publication November 2, 1960. This research was supported in part by Grant No. M-3461(A) from the National Institutes of Health, Public Health Service, and was done at the Walter Reed Army Institute of Research.

[1] Charles Wheatstone, Contributions to the physiology of vision: II. On some remarkable, and hitherto unobserved, phenomena of binocular vision, *Phil. Trans. roy. Soc. (London)*, Part I, 142, 1852, 1-18.

[2] C. H. Judd, Some facts of binocular vision, *Psychol. Rev.*, 4, 1897, 374-389; O. S. Adams, Stereogram decentration and stereo-base as factors influencing the apparent size of stereoscopic pictures, this JOURNAL, 68, 1955, 54-68; T. G. Hermans, Visual size constancy as a function of convergence, *J. exp. Psychol.*, 21, 1937, 307-324; T. G. Hermans, The relationship of convergence and elevation changes to judgments of size, *ibid.*, 48, 1954, 204-208; H. A. Swenson, Reported in H. A. Carr, *An Introduction to Space Perception*, 1935, 362-363; Erich von Holst, Die Beteiligung von Konvergenz und Akkommodation an der wahrgenommenen Grössenkonstanz, *Naturwissenschaften*, 42, 1955, 444-445.

[3] Ewald Hering, *Spatial Sense and Movements of the Eye*, 1942, 172-173.

reduced the apparent size of a distant object.[4] Experiments involving comparison of apparent size under binocular, monocular, and artificial pupil conditions also have been performed.[5]

Another attempt to isolate the influence of accommodation and convergence on apparent size was the experiment of Frank.[6] Her *S*s compared successively a target viewed directly with a target at the same distance when fixation was nearer or farther than the target. The target appeared smaller when fixation was nearer than the target. Fixation-distance was also controlled in the experiment of Ohwaki.[7] Targets were compared in simultaneous viewing between a condition of constant fixation at the distance of the standard target and a condition of normal unrestricted fixation. Constant fixation reduced the deviations of the judgments of size predicted from the visual angle.

In the recent experiment of Heinemann, Tulving, and Nachmias, comparisons of apparent size were made among conditions of monocular viewing with natural pupil, monocular viewing plus artificial pupil, and binocular viewing plus artificial pupil.[8] It was found that the judgments of size deviated from predictions based on visual angle in monocular vision, but when a very small artificial pupil was used, the deviations disappeared. They also found that the deviations from predictions based on visual angle were present under binocular stimulation with accommodation held constant by means of atropine.

An interpretation of the results of the earlier experiments concerning the influence of accommodation and convergence on apparent size is difficult. Stereoscopic devices have certain limitations. The physical target distance and hence the stimulus to accommodation is usually constant, while the stimulus to convergence is usually varied over a wide range. Much evidence has shown that accommodation and convergence are fairly closely associated with each other under normal viewing conditions.[9] Thus, if the stimuli for accommodation and convergence are made quite incongruous in a stereoscope, the *S* may respond to the one or to the other and report either double images or blurred images. Either result makes the data about apparent size ambiguous, especially where no direct measures of accommodation or convergence are obtained.

The experiments utilizing a fixation-point not at the target-distance also suffer from a lack of measures to verify the distance of actual fixation. In addition, the use of a target not in focus on the retina adds the complication of an unspecified change in size of the retinal image.

In evaluating the influence of accommodation on apparent size, it is necessary to specify quantitatively the size of the retinal image. Accommodation results in a very slight enlargement of the retinal image.[10] This is due to the slight forward movement of the nodal point. The phenomenon we are considering is in the opposite direction; namely, a decrease in apparent size with near fixation. Furthermore, the change in apparent size with fixation-distance is many times larger than the slight change in retinal image-size with accommodation. The situation becomes different, however, when an artificial pupil is added.[11] In this case, accommodation may result in either appreciable enlargement or reduction of the retinal image. This factor has not always been taken into account.

---

[4] N. R. F. Maier, The illusion of size in pin-hole vision, this JOURNAL, 41, 1929, 291-295.

[5] T. G. Hermans, The perception of size in binocular, monocular, and pinhole vision, *J. exp. Psychol.,* 27, 1940, 203-207; A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37.

[6] Helene Frank, Ueber den Einfluss inadäquater Konvergenz und Akkommodation auf die Sehgrösse, *Psychol. Forsch.,* 13, 1930, 135-144.

[7] Sonoko Ohwaki, On the role of accommodation in the size-constancy phenomenon, *Tohoku psychol. Folia,* 14, 1954, 17-31.

[8] E. G. Heinemann, Endel Tulving, and Jacob Nachmias, The effect of oculomotor adjustments on apparent size, this JOURNAL, 72, 1959, 32-45.

[9] I. M. Borish, *Clinical Refraction,* 2nd ed., 1954, 98.

Several experiments have shown that another important variable is the attitude of the *S*. Joynson has explained how 'apparent size' may have two quite different meanings for the experimental *S*.[12] One of these attitudes has been called constancy, objective or judgmental; the other has been called visual angle, retinal, phenomenal, projective, or analytic. Several experiments have shown empirical differences in size-matches as a function of such instructions or attitudes for the *S*.[13] Weber and Bicknell showed a change in size judgments as a function of practice in a stereoscopic situation.[14] Jenkin and Hyman recently showed that the two attitudes are statistically independent.[15] Possibly the difference between the size judgments under these two attitudes depends upon whether *S* takes the different distances of the targets into account. In a completely effective reduction-situation, where no distance-cues are available to the *S*, then *S*'s attitude may make no difference to his size-judgments.[16] But where even slight distance cues are available, the effect of *S*'s attitude may be quite important.

The present experiment undertook to obtain actual measures of accommodation, convergence, and pupil-size at the same time that the judgments of size were obtained. The size-judgments were obtained in a dark room, with cues to distance greatly reduced.

Method. Three experimental conditions involving a progressive reduction of accommodation and convergence were employed to investigate the effect of these reductions upon the judgments of apparent size. These conditions were: (1) binocular viewing, (2) monocular viewing, and (3) monocular viewing through a very small artificial pupil. It was expected that accommodation and convergence would change with distance to the targets under binocular and monocular viewing, and that the artificial pupil would minimize any changes in accommodation and convergence.

---

[10] J. I. Pascal, Effect of accommodation on the retinal image, *Brit. J. Ophthal.*, 36, 1952, 676-678.

[11] Hermann Helmholtz, *Physiological Optics,* I, 1924, 127-128.

[12] R. B. Joynson, The problem of size and distance, *Quart. J. exp. Psychol.*, 1, 1949, 119-135.

[13] B. E. Holaday, Die Grössenkonstanz der Sehdinge bei Variation der inneren und ausseren Wahrnehmungsbedingungen, *Arch. ges. Psychol.*, 88, 1933, 419-486; T. M. Martin and R. W. Pickford, The effect of veiling glare on apparent size relations, *Brit. J. Psychol.*, 29, 1938, 92-103; A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192.

[14] C. O. Weber and Natalie Bicknell, The size-constancy phenomenon in stereoscopic space, this JOURNAL, 47, 1935, 436-448.

[15] Noel Jenkin and Ray Hyman, Attitude and distance-estimation as variables in size-matching, this JOURNAL, 72, 1959, 68-76.

[16] A. H. Hastorf and K. S. Way, Apparent size with and without distance cues, this JOURNAL, 47, 1952, 181-188.

The factor of the attitude or set of the *S* was controlled as much as possible by instructions. Both analytic and objective instructions were employed. They were tested in the following sequence: (1) analytic, (2) objective, and (3) analytic. Additional information about the absolute sizes and distances of the targets was supplied to *S*s during the size-matching sessions under the objective instructions. There were two purposes in thus varying the instructions and information to the *S*s. One was to see if there were any differences either in the judgments of size or the oculomotor adjustments under the analytic and objective instructions. The second objective was to see if the judgments of size or oculomotor adjustments under the analytic instructions were changed as a result of the intervening experience and the additional information supplied during the objective sessions. The three conditions of viewing, binocular, monocular, and monocular plus artificial pupil, were tested under each of the three instructions.

Apparatus. The experiment was performed in a darkened room approximately 7 m. in length. Each target was a diamond-shaped milk-glass surface illuminated from behind with a small 6-v. bulb. This was entirely enclosed. The brightness of the target was 0.1 ft.-lam. at the eye-position. The luminance was increased for the condition using an artificial pupil to provide equal apparent brightness as compared with the targets viewed through the half-silvered mirror.

Fig. 1 illustrates the standard- and comparison-targets in elevation. M_1 is a half-silvered mirror used for the measurements described below. T_1, T_2, T_3, T_4, and T_5 are the standard targets, all of 52′ of visual angle on a side. They were presented at distances of 100, 50, 33, 25 and 20 cm. from the left eye.[17] They were all on the same line of sight for the left eye because of the requirements of the Purkinje-image photography to be described below. Targets T_1, T_2, and T_3 were hinged to be rotated in and out of the line of sight. M_2 and M_3 are half-silvered mirrors to present the two closest targets to the eye. V is the variable comparison-target at 580 cm. from the eye. Its size was controlled by a motor operated by *S*, and a selsyn system provided a remote indication of its size to *E*.

The apparatus was also designed to obtain direct measures of accommodation, convergence, and pupil-size at the same time that *S* gave his judgments of size. To measure accommodation and pupil-size, photographs were made of the left eye. To do this and to have the targets in view at the same time, the photographic apparatus was mounted above the *S* in the vertical plane. A first-surface 12.5 × 17.5 cm. half-silvered mirror M_1, positioned in front of *S*'s eyes and tilted 45° to the horizontal, reflected the view of *S*'s eyes into the recording apparatus. The stimulus-targets could be seen through the half-silvered mirror.

Photographic recording. Accommodation was measured objectively by changes in the size of the Purkinje image reflected by the front of the lens of the *S*'s left eye. Fig. 2 presents at the left a schematic view of the apparatus as seen facing *S*. A xenon flash bulb F, a double slit S_1 and a lens L_1 situated on the left of the visual axis, provided the light-source for the Purkinje images. This was reflected in the first-surface half-silvered mirror M_1 (not shown in the front view) to the left eye, *O.S.* The Purkinje images formed by the surfaces of the lens were photographed by the camera C situated on the right of the visual axis. The Purkinje images were calibrated against a subjective coincidence optometer P, shown in both the front and side views of the apparatus. The stimulus-targets were placed in a line in front of the left eye of *S* so that only the right eye changed its position to converge on the targets. Further details on the Purkinje-image photography and calibration are presented elsewhere.[18]

---

[17] Throughout this paper, the reference-point for all angles and distances is the center of rotation of the eye, assumed to be 13 mm. behind the cornea.

When the artificial pupil was used, the half-silvered mirror was replaced by a first-surface fully silvered mirror at the same position. The silvering was removed to produce an artificial pupil 0.3 mm. in diameter on the line of sight of the left eye. The surface of the mirror was 35 mm. in front of the left cornea.

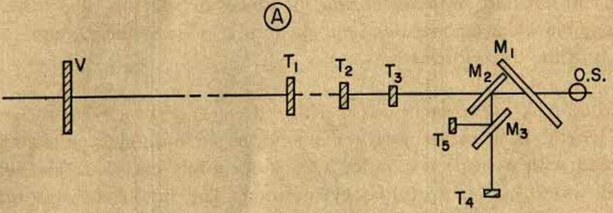


FIG. 1. SCHEMATIC DRAWING OF THE VISUAL TARGET ASSEMBLY
Side view, not to scale.

The photographs of the left eye also provided measures of the diameter of the pupil. This was calibrated by photographing a small millimeter rule in the same plane as the pupil. *S*'s head was held in position by a dental impression bite and a forehead rest.

Haploscope. Visible to *S* was a vernier target consisting of one vertical line of light to each eye. The optometer P in Fig. 2 provided a line of light to the left eye. The other line of light came from an illuminated target H rotating about the center of rotation of the right eye. This is shown in both views in Fig. 2. The apparatus was adjusted to match the interpupillary distance of each *S*. The line for the right eye was presented slightly below the horizontal plane. In this position, the stimulus-targets could be occluded for the right eye, and the haploscope line was still visible. The two haploscope lights appeared about one degree to the right of the size target. To measure convergence, *E* rotated the haploscope target H until *S* reported coincidence of the lines while fixating on the size-target.

Procedure. The experiment was conducted in a darkened room. Only one target was illuminated at a time. *S* viewed first a near standard and next the far comparison-target. When the comparison-target was illuminated, *S* adjusted its size. Standard- and comparison-targets were presented alternately until *S* was satisfied with their equality. Flash photographs of pupil-size and accommodation were taken on the first three viewings of each target. Convergence was measured on the first viewing of each target, and then the haploscope lights were turned off.

---

[18] W. R. Biersdorf, The utility of the anterior lens Purkinje image as a measure of accommodation, *Amer. J. Optom.*, 37, 1960, 352-362.

Within each experimental session, each standard target was matched twice, once with the starting position of the comparison-target too large, once too small. Thus, with 5 standards, 10 size-matches in each session were presented in random order.

Each *S* was given 12 sessions with the analytic instructions first, and attempts were made to keep him naïve concerning the distances and true sizes of the targets.

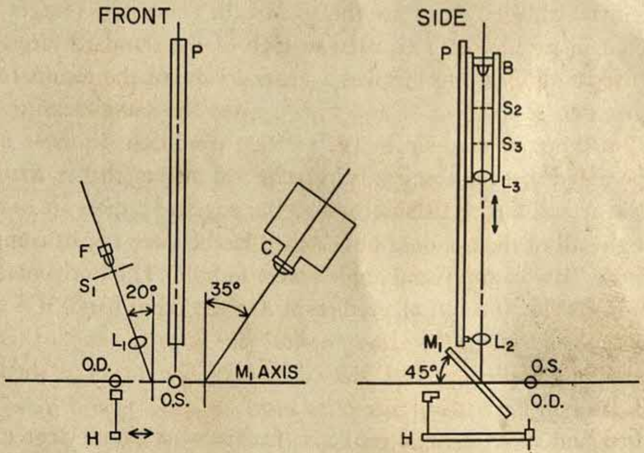


FIG. 2. SCHEMATIC DRAWING OF THE OCULOMOTOR MEASUREMENT APPARATUS

Four sessions were conducted with each of the three viewing conditions in the order, artificial pupil first, followed by the monocular, and finally the binocular. Then *S* was given two sessions under the objective instructions with each condition: binocular, monocular, and monocular-artificial pupil, in that order. Finally, *S* was returned to the analytic instructions for two sessions on each of the conditions in the original sequence.

Under the analytic instructions, *S* was asked to disregard the 'real' size of the targets and to disregard the distances of the targets. He was told to make the targets look the same angular size. This was explained by asking him to imagine taking a photograph of each of the targets, and then placing one photograph directly on top of the other and making the area covered by the targets the same. An effort was made to supply as little information as possible about the objective set-up.

Under the objective instructions, *S* was asked to make the far comparison target the same 'real' or physical size as each near standard. He was told that if the near target appeared to be one-half inch on a side, he should set the far target also to be one-half inch on a side. For sessions with these instructions, *S* was also given additional information about the targets. At the beginning of each session, *S* was allowed to stand to one side of the experimental position and see all the targets illuminated individually. The standard targets were verbally numbered one through five and he was informed that he would be told which of the five standard targets or the comparison-target he was viewing at all times.

Three *S*s, one man and two women, with 20/20 vision or better as tested by a Snellen chart at 20 ft., were used. They did not wear glasses and ranged in age from 20 to 30 yr.

Results: Size-judgments. The data for all three *S*s are contained in Table I, in which the entries are the size of the comparison-target (length of diagonal in millimeters) equated to each of the standard targets under each of the conditions. Fig. 3 gives a general view of the results for one *S* in the first two series; Fig. 4 and Fig. 5 gives the same measures for all three *S*s, in this case grouped by the viewing condition. In these figures it should be noted that the abscissa is in reciprocal meters, that is, that '5' represents the nearest target. Inasmuch as all the standard targets are of constant visual angle, all of them should be matched by the same size of comparison-target if the "law of the visual angle" were to hold. The horizontal dashed line represents this theoretical prediction. On the other hand, if *S* matched the comparison-target to the same physical size as each standard target, the result would be called perfect size-constancy. The lower dashed line is the prediction on the basis of size-constancy.

The first and most obvious result is that there is a high degree of constancy under the objective instructions and very much less constancy under the analytic instructions. Moreover, under the latter instructions, the particular viewing conditions is important, binocular versus monocular versus artificial pupil. All other effects are secondary to these basic effects.

Binocular vs. monocular. It will be recalled that the *S*s served under the analytic condition first. Accordingly, a comparison of the three viewing conditions affords a test of the influence of convergence and accommodation relatively uncontaminated by knowledge of the *S*.

The binocular data in Fig. 3 for S_Z show the expected deviation from the "law of visual angle." The nearer standards were matched more closely by a smaller comparison-target than were the more distant ones. The monocular data showed a similar deviation in that the nearer targets were also matched to a smaller comparison-target. The absolute monocular sizes, however, were larger than the binocular sizes for all five standard targets. A similar binocular-monocular relation also held for S_B. The S_S showed the decreasing size for the nearer standard targets with no differences be-

TABLE I

Size of the Comparison-Target (in mm.) for the Different Target-Distances

Instructions	*S*	Binocular					Monocular					Artificial pupil				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Analytic 1	Z	88	80	70	55	50	110	96	90	74	63	123	138	151	159	183
	B	106	92	76	66	63	116	110	93	81	79	128	134	149	150	152
	S	116	112	108	96	99	118	112	108	107	104	126	134	150	161	185
Objective	Z	32	22	19	12	6	26	20	16	9	4	36	25	19	12	8
	B	53	39	28	17	12	45	33	29	19	13	57	46	44	37	32
	S	12	8	6	3	3	14	8	6	5	5	20	15	12	11	11
Analytic 2	Z	72	56	43	34	23	96	75	62	48	48	109	130	143	143	180
	B	102	88	76	62	57	102	93	75	64	59	117	132	139	140	136
	S	113	112	106	88	97	118	107	108	99	98	122	112	140	150	166
												127	—	—	151	190
											Simultaneous	128	—	—	154	181
												129	—	—	156	191

tween the binocular and monocular data. Thus, both the binocular and monocular data showed that when accommodation and convergence were changing normally as a function of target-distance, the size-judgments showed a systematic deviation from the law of the visual angle.

Artificial pupil. The data using the artificial pupil must be referred to a different standard than the binocular and monocular data. The very small artificial pupil enables S to see all the squares clearly, regardless of his state of accommodation. As will be shown, S_Z maintained her accommodation constant at a relatively far position. When she looks at the far target, the size of the retinal image is relatively unaffected by the artificial pupil. When viewing the near target with the far accommodation, however, the retinal image is enlarged, as pointed out by Helmholtz.[19] Thus, the prediction for the experimental data was that the size-matches by the comparison-target would become larger as the standard targets were closer to the S. The data for S_Z fulfilled the condition of unchanged accommodation at the farther distance and the top data line in Fig. 3 showed the predicted enlargement of size-matches for the nearer standards.

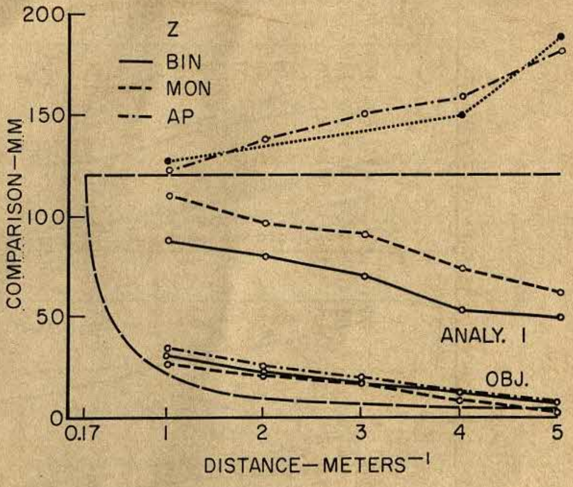


FIG. 3. SIZE-MATCHES FOR S_Z UNDER INITIAL ANALYTIC AND OBJECTIVE INSTRUCTIONS: BINOCULAR, MONOCULAR, AND ARTIFICIAL PUPIL VIEWING

Abscissa for this and all subsequent figures is the target-distance in reciprocal meters. Ordinate is the diagonal of the comparison square in millimeters.

---

[19] Helmholtz, *op. cit.*

Under the opposite situation, when accommodation is maintained constant at a very near distance while fixating a far target, the retinal image of the far target is reduced below what it would be if the artificial pupil were not present.[20] But the predicted result is the same, since the comparison-target must be made larger to match the near targets. S_S maintained his accommodation relatively constant at a near distance, and his judgments of size are similar to those of S_Z. This can be seen in Table I and also in the top part of Fig. 5, where the data for the three *S*s are presented.

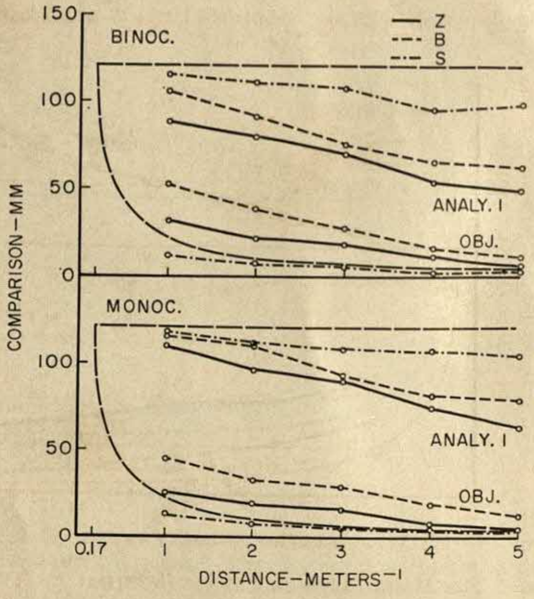


FIG. 4. SIZE MATCHES FOR ALL *S*S UNDER INITIAL ANALYTIC AND OBJECTIVE INSTRUCTIONS: BINOCULAR AND MONOCULAR VIEWING
Ordinates same as in Fig. 3.

The data for S_B, shown in Fig. 5, deviated from those obtained from the other two *S*s for Targets 4 and 5. This can be attributed to this *S*'s change in accommodation among targets which occurred even when using the artificial pupil. This was particularly true for Target 5, as can be seen in Table III. When accommodation is not held constant, the size of the retinal image is less predictable.

---

[20] Helmholtz, *op. cit.*

Control experiment. The amount of change in the size of the retinal image will vary with the distance of the artificial pupil from the eye. The amount was quite large in this experiment, due to the appreciable distance between the eye and mirror containing the artificial pupil.

To obtain a quantitative prediction of the amount of enlargement in the retinal image, a control experiment was added. After the sessions with the artificial pupil under the initial analytic instructions, *S* viewed *simultaneously* a near standard and the far comparison-target through the artificial pupil. Targets 4 and 5 could easily be viewed in this manner, and by moving Target 1 slightly to one side, it could also be seen simultaneously with the comparison-target. *S* was asked to equate the sizes. The results for Squares 1, 4, and 5 are found in Table I at the bottom, labeled "Simult.," and the values for S_Z are also shown in Fig. 3, as the dotted line at the top. As can be seen, the experimental values with successive viewing were quite similar to the control values found with simultaneous viewing.

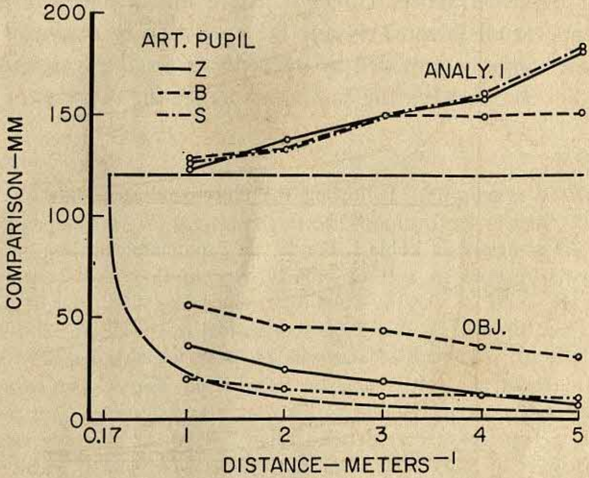


FIG. 5. SIZE-MATCHES FOR ALL *Ss* UNDER INITIAL ANALYTIC AND OBJECTIVE INSTRUCTIONS: ARTIFICIAL PUPIL VIEWING

Ordinates same as in Fig. 3.

Objective instructions. To this point we have been considering results under the initial analytic instructions. We turn now to the results for objective instructions; the judgments of S_Z are illustrated as the bottom three lines in Fig. 3. Mode of viewing (binocular, monocular, artificial pupil) apparently made no difference with her. Small differences appeared with mode of viewing for the other two *Ss*, but the values in no case approached the extreme differences found under the analytic instructions.

Fig. 4 shows what happens when the instructions are changed under binocular and monocular conditions for all three *S*s. The upper three lines represent the judgments of size under the initial analytic instructions, while the lower three lines represent the judgments under objective instructions. The objective judgments fall close to the line representing constancy both for binocular and monocular viewing. There is no overlap in either graph for the two instructions.

The data for the condition employing the artificial pupil are illustrated in Fig. 5. Under the analytic instructions, the size-judgments showed the enlargement discussed above. Under objective instructions, however, the size-judgments for all *S*s came close to the prediction of *constancy* in spite of the optical handicap provided by the artificial pupil. Again there is no overlap in the data between the judgments under the two sets of instructions.

Second analytic instructions. Following the objective instructions, the *S*s were returned to the analytic instructions. The results were quite variable among the *S*s. All the data are contained in Table I. For S_Z, the oculomotor adjustments remained unaffected by instructions as will be seen below, but the size-judgments showed changes. For all modes of viewing (binocular, monocular, and artificial pupil) the judgments of size became smaller. Thus they shifted in the direction of constancy.

The other two *S*s' judgments of size with binocular viewing remained similar to those obtained under the initial analytic instructions. For S_B her second set of monocular judgments became similar to her binocular judgments. This was accompanied by changes in convergence detailed below toward her binocular measures. S_S showed no changes in monocular judgments. These were similar to his binocular judgments in both sets of analytic instructions.

For viewing through the artificial pupil, S_B and S_S also showed changes in size-judgments between the two sessions under the analytic instruction. Both *S*s began changing their accommodation and convergence during the intervening objective sessions. These oculomotor changes persisted under the second analytic instructions and were accompanied by differences in apparent size from that reported under the original analytic instructions.

S_S showed no oculomotor changes in the first session under the second analytic instructions and large changes in the second session. His judgments of size on the first session were exactly the same as his judgments under the original analytic instructions. On the second session, the judgments of size showed reductions in the direction of constancy. Thus when accommodation was constant, whether at near or far, the judgments of size equalled a prediction based on retinal size. When accommodation varied, the combination of optical and perceptual factors changed the resulting judgments of size.

Oculomotor adjustments. The results of the measurements of convergence and accommodation are presented in Tables II and III, and the more significant comparisons are shown in Fig. 6 and Fig. 7.

In Table II, the measurements of convergence for the three *Ss* under all conditions are presented. The measures are in degrees deviation of the line of sight of the right eye from its straight forward position. Each entry is the over-all mean of 2 values for each standard target and 10 values for the comparison-target within each session. The measures may be converted to the distance of fixation by means of each *S*'s interpupillary distance, 60 mm. for S_Z, 63 mm. for S_B, and 65 mm. for S_S.

Table III presents the measurements of accommodation for the three *Ss*. Each entry is the overall mean of 6 values for each standard target and 30 values for the comparison-target within each session. The measures are in millimeters from the reader used to measure the photographic records. The units are arbitrary and depend on the adjustments for each *S*. Thus, the values cannot be compared across *Ss*.

Mean values of the diameter of the pupil were based on the same number of readings as the values for accommodation. Pupil diameter did not change appreciably and often no significant changes occurred until fixation was as close as Targets 4 and 5.

Conditions of viewing. Typical results under the three conditions of viewing, binocular, monocular, and artificial pupil, are shown for S_Z in Fig. 6. These results are only for the first analytic instructions but are typical of all sessions for this *S*. The top section shows the convergence. For binocular viewing, this *S*'s convergence increases nearly linearly as the reciprocal of the viewing distance. The slight departures for distance 5 may have been due to an obstruction in the apparatus. For S_B and S_S, the results were more linear.

Compared with the result under the binocular condition, the convergence for this *S* under the monocular condition showed overconvergence for the far distance and underconvergence for distances of 50 cm. and less. Two of the three *Ss* (S_Z and S_B) showed a similar pattern, while the third *S* (S_S) showed no consistent difference between binocular and monocular viewing.

The artificial pupil eliminated any changes in convergence among targets for S_Z. Her eyes were directed as if on a target at 100 cm. distance regardless of the true distance. The other two *Ss* showed slight changes in spite of the artificial pupil. Thus S_B showed a slight increase in convergence for Target 5. This was attributed to her high visual acuity (20/15 for each eye). It seemed likely that because of this high acuity, the closest target stimulated accommodation somewhat in spite of the small diameter of the artificial pupil.[21] S_S showed slight changes in convergence but the convergence was considerable for all target distances. This behavior changed, however, in subsequent conditions after he received more information about the stimulus-situation.

---

[21] Gerald Westheimer, The effect of spectacle lenses and accommodation on the depth of focus of the eye, *Amer. J. Optom.*, 30, 1953, 513-519.

TABLE II

Convergence (in Degrees) for the Target-Distances in Reciprocal Meters

Instructions	*S*	Binocular						Monocular						Artificial pupil					
		0.17	1	2	3	4	5	0.17	1	2	3	4	5	0.17	1	2	3	4	5
Analytic 1	Z	1.7	4.1	7.2	10.0	13.4	14.4	3.8	4.0	5.6	8.2	11.8	13.5	4.0	4.1	3.9	4.2	3.6	3.8
	B	2.0	4.9	8.2	11.8	14.8	17.8	3.2	3.8	6.4	9.3	12.2	14.4	3.6	3.7	3.5	4.0	4.0	5.3
	S	1.6	4.3	7.7	11.6	15.0	18.2	1.7	4.0	8.0	12.0	16.3	19.4	14.6	14.1	14.6	15.6	15.8	17.6
Objective	Z	1.5	4.0	7.2	11.2	14.1	15.2	2.8	3.8	5.2	7.6	10.7	12.8	4.1	4.4	4.1	3.6	3.6	4.5
	B	2.3	5.0	8.1	12.1	14.8	17.8	3.2	5.1	7.3	9.7	14.5	15.6	3.5	4.0	4.1	5.0	6.0	7.6
	S	1.6	4.2	7.6	11.6	15.0	18.4	1.4	3.4	6.4	10.3	14.8	18.2	2.1	6.2	7.6	7.4	12.0	14.8
Analytic 2	Z	1.5	3.9	7.6	10.9	13.8	14.6	2.9	3.3	5.2	7.7	10.8	12.3	3.6	3.2	3.4	3.1	3.1	3.2
	B	2.0	4.8	8.1	11.8	14.6	17.6	3.6	5.0	7.5	11.8	14.4	15.2	3.2	3.7	3.2	4.1	5.4	6.6
	S	1.7	4.7	7.7	11.9	15.1	18.4	1.4	4.3	7.2	11.6	14.6	19.0	3.2	5.4	6.6	6.6	6.0	8.3

TABLE III

Accommodation (in MM.) for the Target-Distances in Reciprocal Meters

Instructions	*S*	Binocular						Monocular						Artificial pupil					
		0.17	1	2	3	4	5	0.17	1	2	3	4	5	0.17	1	2	3	4	5
Analytic 1	Z	8.20	7.18	6.16	5.34	4.88	4.49	7.53	7.20	6.40	5.14	4.85	4.22	6.70	6.85	6.64	6.70	6.60	6.78
	B	8.92	8.12	7.30	6.38	6.08	5.02	8.11	7.96	6.98	6.08	5.67	5.06	7.95	8.02	8.02	7.86	7.78	7.26
	S	8.19	8.03	7.30	6.89	5.80	5.35	8.19	8.08	7.41	6.52	5.90	5.13	5.82	5.59	5.51	5.62	5.27	5.17
Objective	Z	7.95	6.93	5.93	4.92	4.67	4.25	7.56	7.08	6.07	5.30	4.82	4.25	6.65	6.68	6.64	6.48	6.72	6.57
	B	8.70	8.02	7.10	6.17	5.65	5.52	8.62	8.04	7.06	6.28	5.71	5.25	8.45	8.30	7.62	7.88	6.92	6.53
	S	8.08	7.80	7.30	6.80	5.76	4.27	7.94	7.97	7.58	6.82	6.08	5.37	7.93	7.40	6.93	6.96	6.63	–*
Analytic 2	Z	7.34	6.55	5.82	4.96	4.44	4.02	7.30	6.71	6.22	5.38	4.95	4.51	6.94	6.85	6.98	6.92	6.92	6.96
	B	–*						8.02	7.71	7.08	5.98	5.65	5.32	8.00	8.06	7.85	7.49	7.42	6.55
	S	–*						8.06	7.56	6.76	5.86	5.60	4.83	7.82	7.45	7.08	7.14	6.88	6.22

* Data lost in photographic errors.

The measures of accommodation for S_Z are presented in the middle section of Fig. 6. The ordinate is inverted so that increasing accommodation is upward. The binocular and monocular results showed considerable changes in accommodation as a function of target distance, with not much difference from each other. This was also the case for the other two *S*s.

When S_Z looked through the artificial pupil her accommodation remained constant for a distance between 100 and 50 cm. Thus the accommodation was consistent with the convergence. For S_B, both accommodation and convergence showed slight increases at an intermediate distance. For S_S, both accommodation and convergence showed slight increases at a high level, equivalent to a distance between 20 and 25 cm.

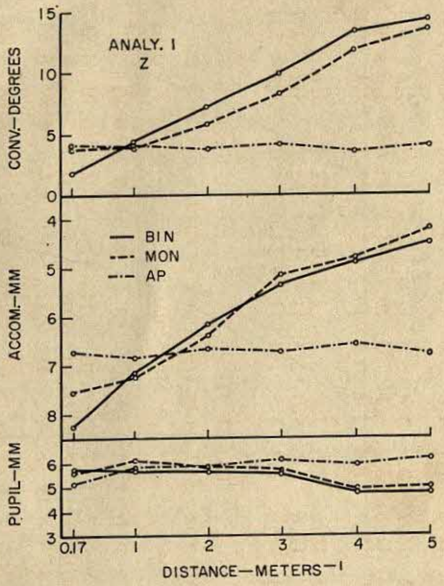


Fig. 6. Oculomotor Adjustments for S_Z under the Initial Analytic Instructions: Binocular, Monocular, and Artificial Pupil Viewing

The lowest section of Fig. 6 shows the effect of viewing conditions on pupil-size. Distance reduced pupil-size only when the target was 25 cm. or less from the eyes. The artificial pupil nearly eliminated this change.

Effect of changed instructions. A comparison of the three types of instructions under binocular viewing is shown for S_Z in Fig. 7. The oculomotor adjustments were not affected by the change from analytic to objective instructions. Table II shows that the small but consistent difference between binocular and monocular convergence was retained, while convergence when using the artificial pupil remained constant.

For all three *S*s, the binocular data did not show differences upon the change to objective instructions. There were tendencies, however, for the measures of two *S*s under monocular conditions and using the artificial pupil to be affected by the size-matching instructions. Monocular convergence-values for S_B tended to approach the binocular convergence-values. Both S_B and S_S showed some changes in the appropriate direction in accommodation and convergence even while using the artificial pupil. Thus, they responded appropriately to their knowledge of the objective conditions and some fraction of this response persisted under the second analytic instructions.

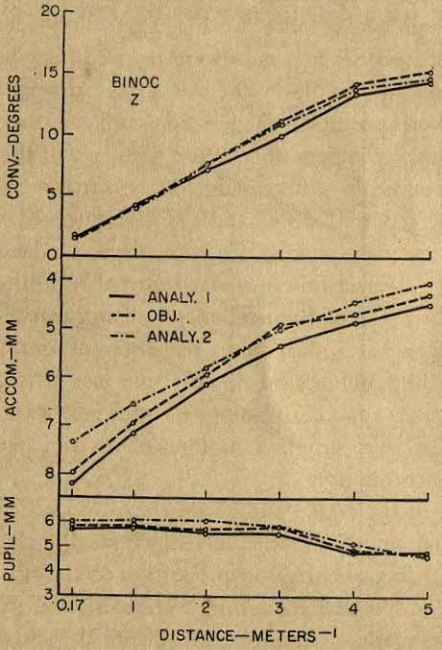


FIG. 7. OCULOMOTOR ADJUSTMENTS FOR S_Z UNDER BINOCULAR VIEWING INITIAL ANALYTIC, OBJECTIVE, AND SECOND ANALYTIC INSTRUCTIONS

Discussion and Conclusions

To summarize the results most generally, it can be said that the oculomotor adjustments were influenced most strongly by the adequacy of the visual stimulus available to *S* and only in a very limited and secondary fashion by changes in the instructions and the knowledge of *S* about the objective conditions. Thus, under binocular conditions of viewing both accommodation and convergence were accurate and adequate, and were not affected by attitude. When cues were reduced, first by restricting *S* to monocular viewing and finally by the use of the artificial pupil, the oculomotor changes were progressively less adequate. At the same time there was a relatively small compensatory effect, with two of the three *S*s, so that changes were produced by a knowledge of the situation.

In contrast, the reported size-changes were influenced to a major extent by the instructions given and by *S*'s knowledge. Under the objective instructions, almost full size-constancy was reported regardless of the viewing conditions and therefore of the oculomotor changes. Although the experiment was conducted in a darkroom, reduction of distance-cues was not complete. Thus, the combination of objective instructions and the added stimulus-information apparently supplied sufficient cues for nearly full size-constancy. The cues supplied by binocular vision were not necessary for reports of size-constancy.

In the absence of the added stimulus-information and objective attitude, as under the initial analytic instructions, much smaller changes in apparent size occurred, and they occurred in proportion to the oculomotor changes. The size-changes in monocular viewing were of similar magnitude to those reported by Heinemann, Tulving, and Nachmias,[22] while the binocular changes in reported size were larger. In general there was a correlation among the three variables of objective distance, the oculomotor adjustments, and the judgments of size, but the correlation was not very exact. The quantitative results presented give a basis for evaluating the order of magnitude of the dependencies involved.

---

[22] Heinemann, Tulving, and Nachmias, *op. cit.*

CONTRIBUTING FACTORS IN THE PERCEPTION OF THE OCULOGRAVIC ILLUSION

By BRANT CLARK, San Jose State College, and ASHTON GRAYBIEL, U. S. Naval School of Aviation Medicine

Perception of the visual horizontal or vertical in normal daily activities is a straightforward and remarkably accurate process which is determined by a perceptual framework in which visual and postural (gravitational) cues covary with minimal discrepancies.[1] In special situations, however, notably in aircraft during turns and in special rotating devices, visual and postural cues may become discrepant and *O*'s perceived horizontal or vertical will reflect this disparity.

Most investigators agree that the perceived horizontal is a function of both visual and gravitational cues acting jointly, but there has been some difference of opinion with regard to the relative importance of each. Gibson has proposed a rapprochement of these opposing points of view and believes that, "The evidence indicates clearly that the apparent visual vertical, as judged by an experimental observer is determined by *both* visual stimulation and postural (gravitational) stimulation acting jointly."[2] The basic question for Gibson then becomes, "How do they interact?" In a recent study Clark and Graybiel found that an antecedent visual framework had a very marked effect on the perceived horizontal in darkness when *O*s were subjected to centrifugal force.[3] These results suggested that it would be worthwhile to make a detailed study of this effect as one aspect of the interaction of visual and postural cues in the perception of the horizontal.

---

* Received for publication July 24, 1961. This study was conducted under the sponsorship of the Office of Life Sciences Programs, National Aeronautics and Space Administration (Grant R-1).

[1] Brant Clark and Ashton Graybiel, Visual perception of the horizontal following exposure to radial acceleration on a centrifuge, *J. comp. physiol. Psychol.*, 44, 1951, 524-534; J. J. Gibson, The relation between visual and postural determinants of the phenomenal vertical, *Psychol. Rev.*, 59, 1952, 370-375; J. J. Gibson and O. H. Mowrer, Determinants of the perceived vertical and horizontal, *Psychol. Rev.*, 45, 1938, 300-323; Kurt Koffka, *Principles of Gestalt Psychology,* 1935, 121-122; C. W. Mann, Visual factors in the perception of verticality, *J. exp. Psychol.*, 44, 1952, 460-464; H. A. Witkin, Perception of the upright when the direction of the force acting on the body is changed, *ibid.*, 40, 1950, 93-106.

[2] Gibson, *op. cit.*, 370-375.

[3] Clark and Graybiel, Visual perception of the horizontal during prolonged exposure to radial acceleration on a centrifuge, *J. exp. Psychol.*, 1962, 63, 294-301.

When *O* is subjected to a gradual increase in rate of rotation while he is seated erect in darkness at some distance from the center of rotation, he perceives a changing direction and magnitude of resultant force acting on his body.[4] He feels that his chair is tilted and that he is firmly pressed against the seat. This force is the resultant of centrifugal acceleration (centripetal force) and gravity which makes an angle φ with the gravitational vertical. If *O* faces in the direction of rotation and views a luminous horizontal line perpendicular to his line of sight under these conditions, he will perceive an apparent rotation of the luminous line about its center and in the same direction as the perceived bodily tilt. For example, if the centrifuge rotates counterclockwise, the luminous line will appear to rotate clockwise. The general nature of this phenomenon has been known since the time of Mach and is a special case of what Graybiel has called the oculogravic illusion.[5] The oculogravic illusion, then, presents a convenient technique by which to study the perception of the horizontal in darkness when both visual and postural cues can be controlled. A line-of-subjective-horizontal (*LSH*) can be readily established under static conditions with a very small error, and deviations from the static *LSH* during rotation can be determined. The purpose of the present paper was to measure changes in the perceived visual horizontal in darkness following varied exposures to a complex visual field while *O* was subjected to a change in magnitude and direction of the resultant force acting on his body.

METHOD

Apparatus. All of the measurements were made in the Pensacola Slow Rotation Room which always rotated counterclockwise with *O* facing in the direction of rotation.[6] A specially constructed seat was mounted firmly near the door 6.25 ft. from the center of rotation of the room (Fig. 1). The seat had three basic parts: (1) a bucket seat with cloth support to brace *O*'s back; (2) a fiberglass helmet molded to fit the head closely and hold it rigidly in position; and (3) a collimated, luminous line of light mounted directly in front of *O*. The luminous line could be rotated about its center by either *E* or *O*. To maintain a clear, visual frame of reference, a white cardboard 20 x 22 in. was mounted 12 in. in front of *O;* black horizontal lines 4 mm. wide and 2 in. apart were drawn on the cardboard to produce a clear reference to the horizontal.

Observers. Five experienced *O*s, including the authors and three volunteers, were tested. Each was thoroughly familiar with the problem under study and had made many other observations during rotation.

---

[4] Clark and Graybiel, *op. cit.,* 525-534; Witkin, *op. cit.,* 93-106.

[5] Ashton Graybiel, The oculogravic illusion, *Arch. Ophthal.,* 48, 1952, 605-615.

[6] Ashton Graybiel, Brant Clark, and J. J. Zarriello, Observations on human subjects living in a "slow rotation room" for periods of two days, *Arch. Neurol.,* 3, 1960, 55-73.

Procedure. Each series of observations began with the room stationary. *E* offset the line from horizontal, and *O* set the line to horizontal. Ten settings were made in the dark, the line always being offset clockwise. The mean deviation of these ten settings from the physical horizontal was computed for each *O* and was considered to be his static *LSH*. Immediately following these 10 settings, the room was accelerated slowly (in approximately 69 sec.) to establish a 12° deviation of resultant force (φ) from the gravitational vertical. *O* perceived himself tilted to his right, and the luminous line was perceived to rotate slowly clockwise. The angular velocity of the room was held constant within narrow limits to maintain φ at 12° throughout the remaining observations. Settings to the perceived horizontal began at least 1 min. after constant velocity was reached to permit the effects of the stimulation of the semicircular canals to be reduced to a negligible value.

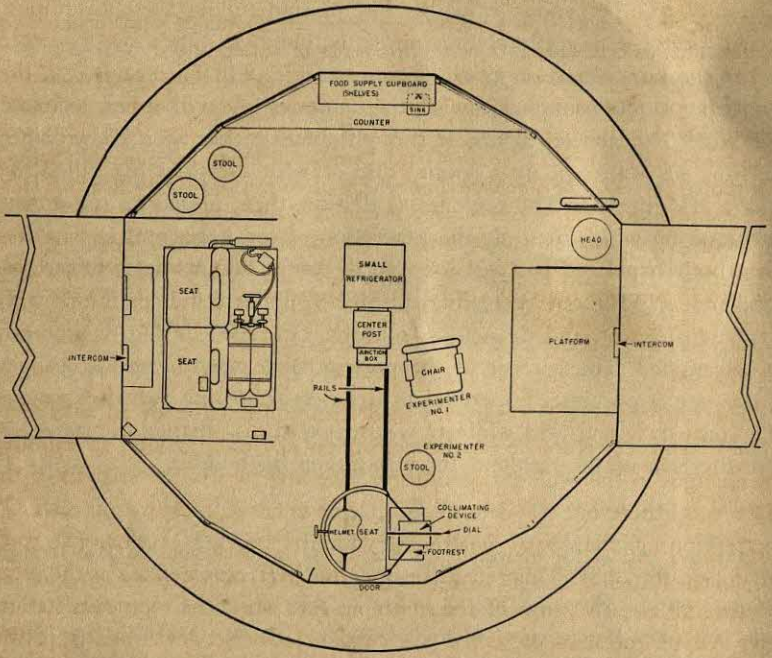


FIG. 1. THE PENSACOLA ROTATION-ROOM

(1) Observations following prolonged visual references. *O* observed the black horizontal lines during rotation with *room lights on for a minimum of 2 min.* before each series of settings began. He was instructed to keep his fixation moving from one point to another continuously during the observational period in order to minimize visual after-images during the observations. The room lights were turned out, and *O* made a series of settings of the line to horizontal, three series of settings being made for each of two experimental methods.

(a) O *setting the line.* This series of observations began with the luminous line set at the static *LSH*. The room lights were turned out, and *O* immediately set the line to horizontal, which in all cases involved rotation of the line counterclockwise. At *O*'s signal that the line was horizontal, one *E* noted the time and a second *E* the setting of the line. *O* then continued to maintain the line at his perceived horizontal, and a succession of settings was made for 5 min., or until *O* had made no change for 1 min.

(b) E *offsetting the line.* This second procedure was identical with the preceding one except that after the first setting was recorded, *E* immediately offset the line clockwise, and *O* reset it to apparent horizontal. These settings were continued for 5 min., and, as before, each *O* made three series of settings.

(2) Observations following brief visual references. All of these observations were made during rotation with a minimum of *2 min. in darkness* before each series began. The luminus line was set at *O*'s *LSH* for darkness. The room lights were then turned on for 1, 5, 15, 30, or 60 sec. During this interval *O* observed the black horizontal lines and the luminous line which, of course, appeared tilted. Again, *O* was instructed to keep his eyes moving to minimize after-images. In setting the line, *O* followed the same procedure used in the first method above. When the room lights were turned out, he immediately set the line to his apparent horizontal, and the two *E*s recorded time and setting. Successive settings were made as before, and three series of observations were made for each of the five exposure-periods.

RESULTS

The purpose of this study was to determine the effect of an antecedent visual frame of reference on the perception of the horizontal as indicated by changes in the oculogravic illusion. The first step in the analysis of the data was, therefore, to determine a representative measure of each *O*'s perception of the visual horizontal with the normal gravitational force acting on the body. This reference line, the *LSH*, was defined as the mean of the initial 10 settings of the luminous line while the room was stationary. All of the data presented are counterclockwise deviations from this line since the perceived displacement was clockwise. The settings during rotation, therefore, give measures of the oculogravic illusion at various points in time following exposure to a clear visual frame of reference to the horizontal. Another significant factor in presenting the results is that the data do not lend themselves readily to conventional statistical analysis. Although time was a critical variable, *O* made his own adjustments, setting his own pace. This resulted in settings which were judged to be horizontal at somewhat different points in time for every *O* and each series. An initial analysis of the data was, therefore, made from plots of setting against time for each *O*'s several series of settings. The data to be presented below, however, are median scores for time and deviation of the setting from the static *LSH*. These were determined as follows: The plots of individual

LIBRARY
J-403
8.9.70

data and the raw data were considered together for the three series for each condition for every *O*. Points as close as possible together in time were identified. Then the median of three scores for time and the median of the three corresponding deviation scores were determined and plotted on a curve for every *O*. Variations in time among these three measures shortly after the lights were turned off were small (2-3 sec.) while after 60 sec., the range of time scores became greater. By this time, however, the major change in the perception had already occurred, and deviations in time were less significant. A comparison of these curves with those of individual series indicates that they give a representative picture of changes in the perception of the horizontal.

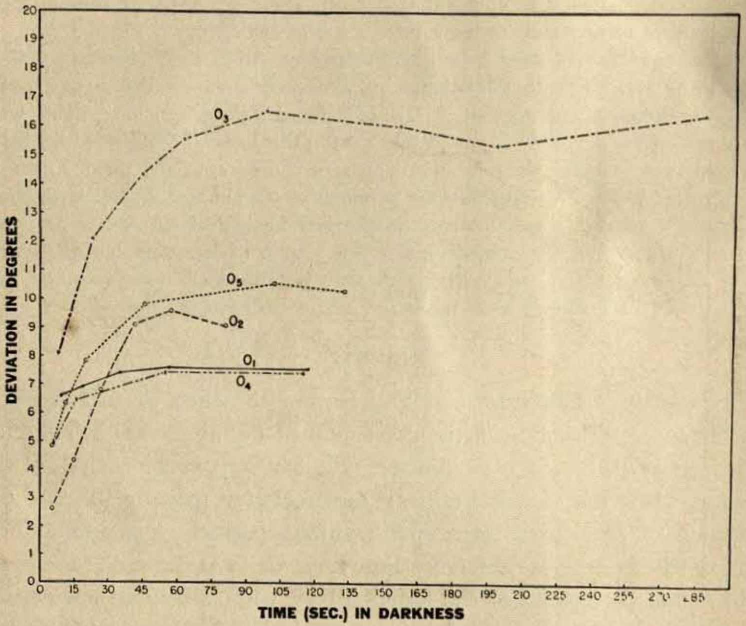


FIG. 2. CHANGE IN PERCEPTION OF THE HORIZONTAL FOLLOWING PROLONGED EXPOSURE TO A VISUAL FRAMEWORK
(*O* set the luminous line continuously.)

(1) Observations following prolonged visual references: (a) O *setting the line continuously.* The median curves for the five *O*s are presented in Fig. 2. These curves terminate at the time of the final settings, but they might be extended at this same value for an additional 60 sec. since *O* continued his observation for 60 sec. without making a change in setting. The general pattern of change in the perception of the horizontal is essen-

tially the same for all *O*s. There is a sharp increase in the oculogravic illusion which continues for 45 to 60 sec., and thereafter the variations are slight. To put it in other terms, the change from a predominantly visual frame of reference to a gravitational frame of reference under this experimental procedure is a rather slow process requiring approximately 1 min. As noted in an earlier report on the same *O*s, there are substantial individual differences in the degree to which individual *O*s adopt the new

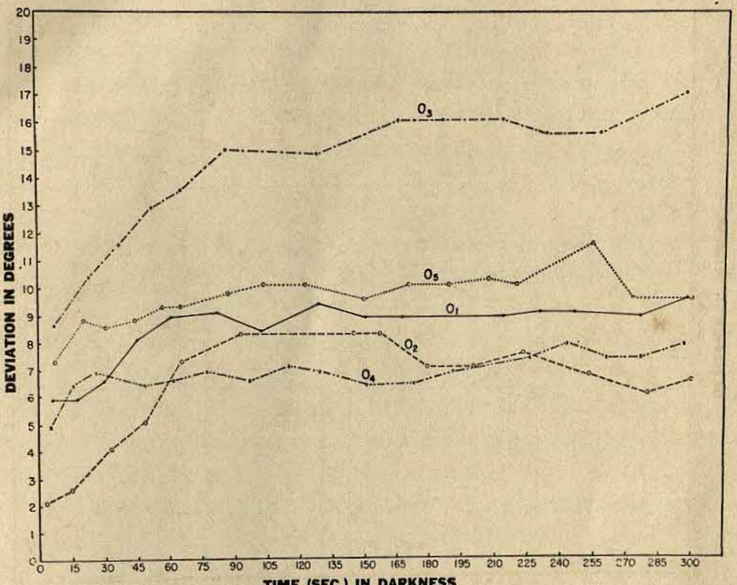


FIG. 3. CHANGE IN PERCEPTION OF HORIZONTAL FOLLOWING PROLONGED EXPOSURE TO A VISUAL FRAMEWORK
(E offset the line.)

direction of resultant force as their frame of reference.[7] These differences, however, do not obscure the fact that the shift from the visual to the gravitational frame of reference requires 60 sec. or more. After this interval of change the oculogravic illusion remains fairly stable up to at least 4 hr.

(b) E *offsetting the line.* Each of these series of observations continued for 5 min., *E* offsetting the line clockwise after the readings for each setting were made. The results parallel the findings where *O* set the line continuously (Fig. 3). The change to the postural cues as the dominant factor in the perception of the horizontal was fairly rapid from 60 to 90

---

[7] Clark and Graybiel, *op. cit.,* 294 f.

sec. Subsequently, there was a gradual increase in the oculogravic illusion up to about 2 min. with a fairly stable condition for the remainder of the 5-min. period. As might be expected, this method produces a somewhat

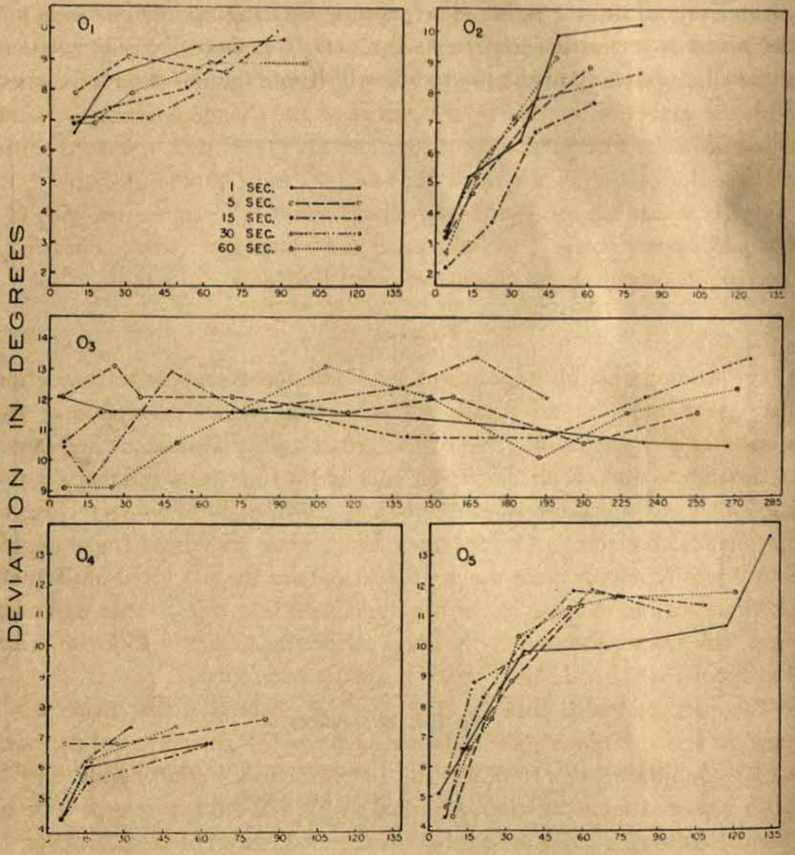


FIG. 4. RESULTS OF THE VARIOUS *O*s

longer period of transition for all of the *O*s. It should be noted also that each *O* tended to maintain his relative position as found in the preceding results (Fig. 2-3). Two *O*s (O_3 and O_5) showed the greatest deviations from the static *LSH* while the other three (O_1, O_2, and O_4) produced curves which overlapped in both methods.

(2) Observations following brief visual references. Here as in the preceding two series of observations the basic data are plots of median values

of time and deviation from the static *LSH* for the five exposure-intervals for each *O* (Fig. 4). In considering these curves, it should be kept in mind again that the period of observation continued for 60 sec. beyond the final setting. An increase in the oculogravic illusion continues for over 60 sec. for all five *O*s, and the terminal settings are very similar to those found in the first part of the experiment. One *O* (O_3), however, shows a change in the oculogravic illusion only to a small degree for the short exposures. With the exception of this *O*, the nature of the change is much the same as found in the first part of the study. Two *O*s (O_2 and O_5) showed shifts of the order of 6-7° while the other two (O_1 and O_4) changed only 2-3° from the initial setting. For these four *O*s and to a lesser degree with O_3, it would appear that a 1-sec. exposure to a visual framework is nearly as effective in orienting the *O*s to the visual horizontal as is 60 sec.

DISCUSSION

The findings in this experiment lend additional support to the notion that visual and postural factors interact in the visual perception of the horizontal and presumably also of the vertical. They also define one aspect of this interaction. With discrepant cues and a full visual reference to the horizontal, *O* immediately perceived the horizontal to be very close to the physical horizontal. On the other hand, when the visual frame of reference was removed, there was a gradual shift in the perceived horizontal; *i.e.* there was an increase in the oculogravic illusion with time in darkness. Thus, the *O*s set the line farther and farther in a counterclockwise direction from the static *LSH* to make it appear horizontal.

The data presented in Figs. 2-4 give the impression that there was a more or less continuous, gradual change in the *O*s' perception of the horizontal. Subjective reports tend to contradict this notion. Only one *O* (O_5) reported some consistent, gradual shifting in the apparent position of the line following exposure to the visual framework. For him, however, as well as for the others, the typical change was a series of rather sudden shifts. More or less suddenly, *O* would become aware that the luminous line was in a new position, and it would be necessary to reset the line using the first method. The changes were similar to the shifts found in observing ambiguous figures. These reports would seem to imply that changing from a visual to a gravitational frame of reference involves a series of steps in reorientation. This is a marked contrast to the reverse condition which ocurs very quickly. These are laboratory examples of experiences widely reported by aviators; *i.e.* there is an accurate, quick orientation in space when adequate visual reference is available, but disorientation following loss of adequate visual cues ordinarily is not immediate but sometimes may appear with shocking suddenness.

These results indicate that this particular variation of the oculogravic illusion serves as a satisfactory indicator of the relationship of visual and gravitational cues to the horizontal. It is also noteworthy that the two methods used give comparable results (Figs. 2-3) and that somewhat greater time is required to reach a stable level of perception when *E* offsets the line. The latter may be interpreted as an error of habituation when *O* set the line continuously. One additional aspect of the method needs comment. The very nature of the method made the *O*s susceptible to after-images of the horizontal lines which might give a continuing frame of reference to the horizontal. The *O*s did report transitory after-effects, but it would appear that keeping the eyes moving was effective in minimizing them. Whatever after-images did occur were vague and short-lived and gave no clear reference to the horizontal. Whatever influence they may have had would appear to have been dissipated within a very few seconds while the major changes in the perception of the horizontal continued far beyond this period of time.

Differences among the five *O*s are marked, both in the slope of the curves and in the terminal level of the oculogravic illusion, the latter varying from 7° to 16°. The median value of the illusion is approximately 9° or three quarters of φ. It seems clear that the nonvisual frames of reference are very different among the *O*s. The typical setting is between the static *LSH* and the direction of resultant force, but closer to the latter. On the other hand, one *O* accepted the gravitational cues to such a degree that the deviation from the static *LSH* was greater than φ. Why all *O*s did not completely accept the gravitational frame of reference and set the luminous line at the value of φ is not clear from these data. One possible explanation would be that there are differences in some egocentric frame of reference to the static, visual horizontal. It also appears to be possible that there may remain some discrepant cues from various proprioceptive mechanisms as well as tactual cues.

SUMMARY

The purpose of this investigation was to study the effect of an antecedent visual frame of reference on the perception of the horizontal when the direction and magnitude of the gravitational force acting on the body were changed. Initial static measures of the perceived horizontal were made with the experimental room stationary. Deviations of the perception of the horizontal from these static measures were determined as a function of

time after exposure to a visual framework while the room was rotating to produce a change in magnitude and direction of resultant force acting on the *O*s. In other words, changes in the amount of the oculogravic illusion served as the criterion of change in the perception of the horizontal. The experiment consisted of two parts: observations following prolonged exposure to a visual framework and observations following varied short exposures. The results showed that the deviation of the perceived horizontal from the static line-of-subjective-horizontal increased fairly rapidly up to about 60 sec. and more slowly for an equal time thereafter. This was found to be true of both long and short exposures. No consistent, systematic differences were observed between exposure time from 1 sec. to 2 min. The results show a gradual, though not a simple, continuous change in the frame of reference from visual to gravitational cues. This is in sharp contrast with the change from a gravitational to a visual framework which occurs very quickly.

SOME RELATIONSHIPS BETWEEN COMPARATIVE JUDGMENT, CONFIDENCE, AND DECISION-TIME IN WEIGHT-LIFTING

By ROSEMARY PIERREL, Brown University, and CATHERINE SAVINO MURRAY, University of Pennsylvania

In psychophysical experiments in which a difference-limen is obtained, three different response-measures have frequently been studied. In the two-category situation, the *S*'s report of comparative magnitude, *e.g.* "heavier" or "lighter," is generally the principal datum. In addition, a measure of reaction-time has sometimes been employed; in this case the latency between the *S*'s exposure to the stimulus-to-be-judged and the making of the report (decision-time) is recorded. Some investigators have also obtained a second report by requesting the *S* to assess his confidence or doubt in the judgment just previously made.

The measure of decision-time was first applied by Cattell in an experiment in which a difference-limen for light-sensitivity was obtained.[1] Cattell noted an inverse relation between magnitude of decision-time and the magnitudes of the differences being judged, a finding confirmed by later investigators. Utilizing an additional measure, Henmon, Johnson, and Festinger have examined the relationships between decision-time, confidence, and stimulus-differences in psychophysical studies of linear extent, whereas Seward has employed similar measures in a visual test of recognition.[2]

In the present study, the interest is in an examination of the correspondences between the comparative judgments, confidence, and decision-times obtained in a weight-lifting experiment. In this experiment, a larger number of *S*s and relatively more trials were utilized than have been employed in most of the related studies previously done. The other purposes of this study may be stated in the form of four questions:

(1) Are the more accurate *S*s more confident than the less accurate?

---

* Received for publication September 18, 1959. This work was done in the Hollingworth Psychological Laboratory of Barnard College.

[1] J. M. Cattell, The relation of time and space in vision, *Proc. Amer. Ass. Sci.*, 48, 1899, 95.

[2] V. A. C. Henmon, The time of perception as a measure of differences in sensations, *Arch. Phil. Psychol. Sci. Meth.*, 1906, (No. 8), 1-75; Time and accuracy of judgment, *Psychol. Rev.*, 18, 1911, 186-201; D. M. Johnson, Confidence and speed in the two-category judgment, 34, *Arch. Psychol.*, 1939, (No. 241), 1-50; Leon Festinger, Studies in decision time: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus differences, *J. exp. Psychol.*, 32, 1943, 291-306; G. H. Seward, Recognition time as a measure of confidence, *Arch. Psychol.*, 16, 1928, (No. 99), 1-50.

(2) Does decision-time vary when errorless trials are compared with trials on which errors occurred?

(3) Does confidence vary when errorless trials are compared with trials on which errors occurred?

(4) How are decision-time and confidence related to the magnitude of the differences being judged?

APPARATUS AND PROCEDURE

The set of weights was made up of six comparison-stimuli (230, 220, 210, 190, 180, and 170 gm.) and a standard (200 gm.). These were constructed from cylindrical plastic pill bottles (2 oz. capacity). The bottles were $2\frac{15}{16}$ in. in height and 1 in. in diameter. Lead shot was placed in the bottles, together with melted paraffin, producing weights of appropriate magnitude. The completed containers were sprayed with flat black paint.

The weights were presented to *S* on a rotating turntable. The second weight to be presented was placed on a microswitch which was mounted with its leaf slightly higher than the top of the turntable. As this weight was lifted, a Standard Electric Timer (model S-I) was activated; this timer was stopped when *S* spoke her judgment into the microphone of a voice-key. A manually-operated mercury switch held the clock circuit open between trials when the weight was not in position on the microswitch. Stimulus-presentation was signalled by a buzzer. Inter-stimulus and inter-trial intervals were timed by a Kodak ⅕ sec. clock.

Eight women students between the ages of 19-21 yr. served as *Ss* for the experiment. Each *S* was seated at a table and blindfolded. The arm used in lifting was placed on a support and the hand allowed to hang free from the wrist 4½ in. above the turntable. A simple wrist motion sufficed to pick up the weights and the arm itself was not lifted from the support during a given trial.

Instructions. The following instructions were read to *S* at the beginning of the session:

> In this experiment you are to compare sets of weights. Place your lifting arm on the armrest so that your hand always hangs freely above a fixed position on the table. On each of many trials I will place two weights near your hand, first one and then the other. When the buzzer sounds, lift the weight under your hand and then replace it. When the buzzer sounds again, lift the weight under your hand and then replace it. Then tell me whether the second weight is heavier or lighter than the first weight. Only the words "heavier" or "lighter" will be accepted, so even if you think the weights are equal, guess "heavier" or "lighter." Say the words "heavier" or "lighter" loudly and clearly and refrain from saying anything else.
>
> After making each judgment, you are to express your confidence that your judgment was correct in the following terms: 50% means average or median confidence, the amount of confidence one usually has in most of one's judgments; 75% means confidence definitely above average, the amount of confidence you have when you are quite sure you are right; similarly, 25% means definitely below average confidence, the amount you have when you are only guessing; 100% is at the extreme end of the scale of confidence and means extreme certainty, no possibility of error; similarly, 0% is at the other end of the confidence scale and means no confidence in your judgment whatever, you are just as likely to be wrong as right. Thus you are to construct a scale of confidence with 50% in the middle, 25 and 75% on each side of the middle range, and 0 and 100% out at the ends of the scale.[3]

---

[3] The last paragraph of these instructions was taken from Johnson, *op. cit.*, 17.

Four seconds were allowed to elapse between the presentation of the first and second weights of a pair to be judged. The inter-trial interval was 5 sec.

Two *Es* were required to conduct the experiment.[4] The first *E* presented and removed the weights, signalled stimulus-presentation with the buzzer, timed the interstimulus and inter-trial intervals, and recorded the comparative judgments. The second *E* manipulated the mercury switch in the chronoscope circuit, reset the clock between trials, and recorded decision-time and confidence-judgment.

The method of constant stimuli was employed and each of the six comparison-weights *(Co)* was compared with the standard *(St)* on 40 trials. The schedule of presentation was a predetermined random order taken from the Rand Tables.[5] In an effort to distribute possible time-error effects, the standard was presented first on odd numbered trials and last on even numbered trials. The data were corrected before processing to take account of this alternation of *St* and *Co* in the first position. Each *S* was given about five preliminary trials, during which data were not recorded, to familiarize her with the procedure. All 240 trials were carried out in a single session lasting approximately 2 hr. Five-minute rest periods were introduced after Trials 80 and 160.

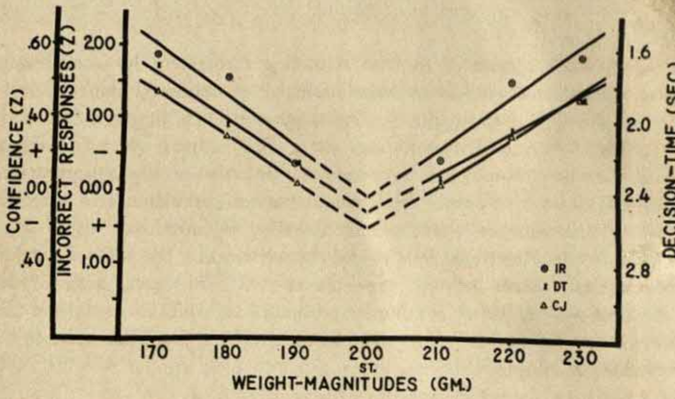


FIG. 1. INCORRECT RESPONSE, CONFIDENCE, AND DECISION-TIME AS A FUNCTION OF COMPARISON-WEIGHT
(Mean data for all the *S*s.)

RESULTS AND DISCUSSION

In Fig. 1 the three measures of response have been plotted as a function of the values of the six comparison-weights. The ordinate of the incorrect responses represents the percentage 'heavier' for *Co* values 170, 180, 190 gm. and percentage 'lighter' for *Co* values 210, 220, 230 gm. The ordinate for the incorrect responses and confidence is in Z-score units. The ordinate for decision-time represents the mean of the values obtained in the experiment. Each point represents the mean value on 40 trials for each of 8 *S*s with each *Co*.

---

[4] Miss Marcia Walder participated in the design and conduct of a pilot study for this experiment. Miss Norma Damashek served as the second *E* for the experiment proper.

[5] Rand Corporation, *A Million Random Digits*, 1955, 1.

As plotted, all three sets of curves tend downward as the magnitude of *Co* approaches *St* and rise as distance from *St* increases. This indicates that, as the magnitude of the difference between *Co* and *St* decreases, there is a corresponding decrease in confidence and an increase in decision-time. Data for all 8 *S*s are given in Table I. Means and sigmas for incorrect response, decision-time, and confidence are presented for each value of the comparison-weight. The means in this table are the basis for the values plotted in Fig. 1.

These observations confirm a number of earlier studies. In Johnson's experiment, both decision-times and confidence were recorded where linear extents were being compared.[6] The data shown in Fig. 1 display relationships similar to those obtained by him. Festinger supported this finding using linear extent, whereas Kellogg demonstrated that decision-time for a light-intensity increases as the differences between stimuli compared becomes less.[7] Volkmann studied confidence and time required for comparative judgments when the inclination of lines was used.[8] He found decision-time to be inversely related to confidence. In Seward's experiment on recognition, novel visual stimuli were intermixed with some stimuli bearing a degree of similarity to those previously presented as well as some stimuli which had been previously shown.[9] Her data indicate that recognition-time is maximal when confidence is minimal.

To examine the degree of relationship between the three measures of response, the Olmsted-Tukey Corner Test of Association was applied to each possible combination of the pairs of variables.[10] When mean incorrect response and confidence are paired for each of the six *Co* values, a significant correlation is obtained. The same result occurs when both incorrect response and decision-time as well as decision-time and confidence are examined. In all of these cases, the measure of correlation was applied to both individual and mean data. The level of significance attained in every comparison was $p < 0.03$.

To show the types of individual data which contributed to the means, two plots are presented in Fig. 2. These points, like those in Fig. 1, represent the three response-measures plotted as functions of the six comparison-

---

[6] Johnson, *op. cit.*, 49.

[7] Festinger, *op. cit.*, 291-306; W. N. Kellogg, The time of judgment in psychometric measures, this JOURNAL, 43, 1931, 65-86.

[8] John Volkmann, The relation of the time of judgment to the certainty of judgment, *Psychol. Bull.*, 31, 1934, 672.

[9] Seward, *op. cit.* 47-48.

[10] F. Wilcoxon, *Some Rapid Approximate Statistical Procedures*, 1949, 10-12.

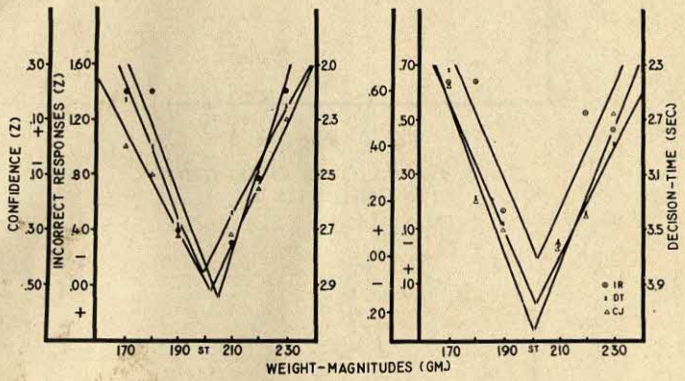
TABLE I

Relative Frequency of Incorrect Response, Confidence Reported, and Decision-Time as a Function of the Comparison-Stimuli

Measure	Comparison-weights											
	170		180		190		210		220		230	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Incorrect (%)	2.5	2.93	6.3	4.78	34.4	5.68	32.9	9.82	6.2	7.75	2.8	2.49
Confidence	71.41	10.21	62.18	11.54	50.55	13.72	51.63	12.42	62.73	14.17	70.39	11.44
Decision-time (sec.)	1.70	0.48	2.02	0.65	2.23	0.66	2.30	0.70	2.05	0.61	1.88	0.50

weights. Each point on each function represents a mean of the 40 trials. Graph A shows the data for the *S* who showed the best correspondence between the three measures of response, whereas Graph B presents the data for the *S* with the greatest variability between measures. As pointed out above, even this *S*'s data (Graph B) yielded a significant association-value when each pair of variables was examined.

The mean confidence and mean decision-times for all *S*s related to each of the *Co* stimuli are given in the plot shown in Fig. 3. The number located near each point represents the magnitude judged on the trials represented by that point. Both decision-time and confidence are systematically related



A B

FIG. 2. INCORRECT RESPONSE, CONFIDENCE, AND DECISION-TIME AS FUNCTIONS OF COMPARISON-WEIGHTS
(Individual data for 2 *S*s.)

to the *absolute* difference between *Co* and *St*. The larger this difference, the greater the degree of confidence and the smaller the value of decision-time. These relations hold regardless of whether the *Co*-stimulus is greater or lesser in magnitude than the *St*.

In the plots shown in Fig. 4, the data have been partitioned in terms of the categories of the confidence-judgment responses, *e.g.* 0, 25, 50%. For each *S*, the mean decision-time was found corresponding to the trials on which each of the given confidence-categories was applied. Each curve represents a single *S*. It is evident from the figure that, although there are large differences among the *S*s in the values of the decision-times corresponding to each category of confidence, there seems to be considerable similarity in the general form of the curve exhibited. *S*-1 and *S*-6 made no confidence-judgments of 0 and 25%, hence these curves have only three

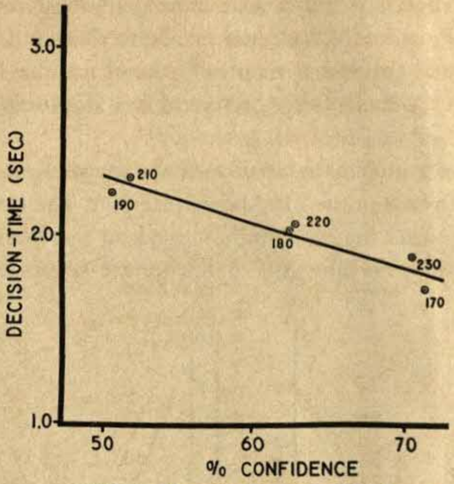


FIG. 3. DECISION-TIME AS RELATED TO CONFIDENCE
(Mean data for 8 *S*s.)

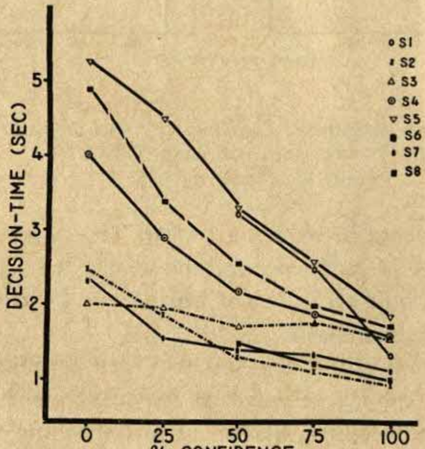


FIG. 4. DECISION-TIME AS RELATED TO CONFIDENCE
(Individual data for all 8 *S*s.)

points. The slightly concave form of the curves is in line with Johnson's observation that as doubt increases arithmetically, decision-time increases geometrically.[11] The marked correspondence between confidence and decision-time suggests that the verbal report of confidence may be an instance in which the *S*'s latency itself might be serving as a discriminative criterion for verbal behavior.

Table II and Fig. 5 offer data bearing on the question of how the confidence expressed relates to the accuracy of the comparative judgments. In Table II, the responses for each *S* have been collected in terms of their confidence. The percentage of judgments that were correct was determined for each category. Finally, the weighted mean was calculated by summing across the *Ss'* percentage correct, weighted by number of judgments in that category, and then dividing by the total use of that category by all *Ss*. By and large, the results are in agreement with Henmon's finding that as there

TABLE II

PERCENTAGE OF CORRECT RESPONSE IN TERMS OF CATEGORIES OF CONFIDENCE EMPLOYED

		Confidence-category				
S		0	25	50	75	100
1	Total use	0	0	59	63	118
	% correct	—	—	79.66	87.30	97.46
2	Total use	36	53	56	51	44
	% correct	61.11	77.36	89.29	96.08	97.73
3	Total use	41	35	63	40	61
	% correct	80.49	80.00	88.89	85.00	98.36
4	Total use	13	80	91	53	3
	% correct	46.15	60.00	90.11	98.11	100.00
5	Total use	40	33	32	59	76
	% correct	70.00	78.79	81.25	89.83	93.42
6	Total use	0	0	70	81	89
	% correct	—	—	74.29	90.12	98.88
7	Total use	45	36	53	37	69
	% correct	57.78	83.33	96.23	97.30	97.10
8	Total use	18	36	60	50	76
	% correct	50.00	69.44	76.67	90.00	90.79
	Grand total	193	273	484	434	536
	Weighted Mean	64.25	72.76	84.71	91.47	96.27

---

[11] Johnson, *op. cit.*, 50.

is an increase in the percentage of correct judgments, *S* responds with higher confidence.[12] This will be seen from the last row of the table. Henmon raised the question of whether the more accurate *S*s tend to use the high end of the confidence-scale, whereas the less accurate *S*s would show a more positively skewed distribution of the judgments of confidence; but we found no support for this conclusion. Similarly, the histograms presented in Fig. 5 offer negligible support for such an hypothesis. This figure shows the relative use of each category by each *S*. The confidence-categories appear from lowest to highest along the abscissa. The eight histograms have been arranged in the order of the *S*s' mean accuracy, the exact value being indicated above each distribution. Our eight *S*s had a rather narrow range of accuracies (90. 42%–79.58%). Only the extreme cases show any suggestion of conformity to Henmon's notion. The two most accurate *S*s tended to use the high categories of confidence most, whereas the least accurate *S* used the 100% category least.

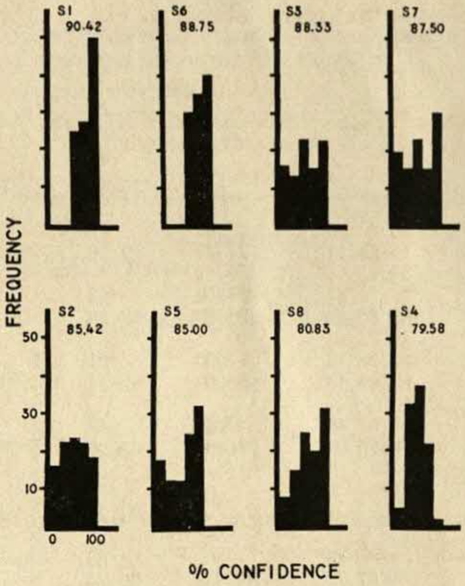


FIG. 5. FREQUENCY-DISTRIBUTION OF CONFIDENCE
(Individual data with *S*s in descending order of accuracy.)

---

[12] Henmon, 1911, *op. cit.*, 201.

An analysis in terms of trials that gave correct and incorrect responses is presented in Table III. The confidence and decision-time is shown for each *S* divided into correct and incorrect trials. In the case of every *S*, the mean decision-time is lower on correct trials than on incorrect trials, and confidence is higher on errorless trials. Tests, (*t*-tests), carried out on the group means given in Table III, show significant differences between error and errorless trials ($p < 0.025$) for both confidence and decision-time. Since the greatest number of errors is made when *Co* approaches *St*, this finding is implicit in the results shown in Fig. 3. These findings with regard to accuracy statistically confirm similar observations previously made by Henmon, Kellog, and Seward.[13]

TABLE III

CONFIDENCE AND DECISION-TIME FOR CORRECT VERSUS INCORRECT RESPONSE

	Confidence		Decision-Time	
S	right	wrong	right	wrong
1	82.83	65.22	2.12	3.00
2	56.10	24.29	1.48	1.93
3	57.11	36.61	1.74	2.06
4	49.75	27.04	2.23	3.33
5	63.85	38.19	3.01	4.15
6	79.22	60.18	1.18	1.42
7	60.48	17.50	1.47	1.89
8	68.04	56.94	2.28	3.31
Group mean	64.67	40.75	1.94	2.64
SD	10.75	16.80	0.55	0.28

SUMMARY

Eight *S*s compared 6 weights with a standard weight in a total of 240 trials for each *S*. The six stimuli used for comparison were symmetrically arranged about the standard. On each trial, three measures were obtained: *S*'s judgment of 'heavier' or 'lighter,' a 5-category statement of confidence (0 to 100%) in the judgment, and the elapsed time between the lifting of the second weight and the voiced judgment.

In the case of both individual and mean data, a significant correlation was found between: (1) confidence and frequency of incorrect response for a given stimulus; (2) decision-time and frequency of incorrect response for a given stimulus; and (3) confidence and decision-time.

As the magnitude of the difference between the comparison- and standard stimuli decreases, there is a corresponding decrease in confidence and an increase in decision-time. The length of decision-time and the degree of confidence are functions, of the magnitude of the difference between the comparison-stimulus and the standard, and does not depend upon which is heavier. A significant difference exists between confidence on correct vs. incorrect trials, and also in decision-times on correct vs. incorrect trials. Correct trials to yield higher ratings of confidence and lower decision-times.

---

[13] Henmon, 1911, *op. cit.*, 201; Kellogg, *op. cit.*, 85; Seward, *op. cit.*, 48.

ATTENTION AND KINESTHETIC AFTER-EFFECTS

By PAUL BAKAN and RICHARD THOMPSON, Michigan State University

Köhler and Adams have recently extended the theory of satiation, originally proposed by Köhler and Wallach,[1] to the relationships between attention and satiation.[2] According to Köhler and Adams, "when attention is concentrated on an object, the underlying cortical process . . . will be intensified, and . . . as a consequence, the object will stand out more clearly; but this intensification of the current will also accelerate its electrotonic action. As a result, satiation in the crucial area will be stronger, and corresponding figural after-effects will be greater than those caused by a comparison-object to which *O* does not attend."[3] They report evidence in support of this hypothesis by showing an effect of attention on both visual and kinesthetic after-effects.

Three experiments bearing on their hypothesis are here reported. The first two represent attempts to replicate their findings; and the third is a test of some of the implications of their hypothesis.

Experiment I. Experiment I is a slight modification of Köhler and Adams' study. We used a tapered wedge to indicate *O*'s point of subjective equality rather than a block arranged in discrete steps, and our calibration of the wedge was in smaller units than the steps of the block.

If attention favors the satiational process which determines the size of the kinesthetic after-effects (*KAE*), then the following results should obtain if *O* has equal exposure to an inspection-stimulus with each hand, while directing his attention only to one hand: If the stimulus-object to which attention is directed is narrower than the standard and is on the side opposite the standard, then the apparent width of the standard should *decrease* after satiation. (2) If it is narrower and on the same side as the standard, then the apparent width of standard should *increase* after satiation. (3) If it is wider and on the side opposite the standard, the apparent

---

* Received for publication July 14, 1961. This research was supported by Grant G-5822 from the National Science Foundation.

[1] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

[2] Wolfgang Köhler and P. A. Adams, Perception and attention, this JOURNAL, 71, 1958, 489-503.

[3] Köhler and Adams, *op. cit.*, 498 f.

width of standard should *increase* after satiation. (4) If it is wider and on the same side as the standard, the apparent width of standard should *decrease* after satiation.

Design. Forty-eight volunteers from the introductory course in psychology served as *O* in this study. They were randomly assigned to one of four experimental groups, each corresponding to the four conditions described above and specified in Table I. The only restriction on random assignment was equalization of sex among the groups. There were 8 men and 4 women in each group.

TABLE I

OUTLINE OF EXPERIMENTAL CONDITIONS FOR EXPERIMENT I
(Standard Stimulus 1.5 in. in width)

Group	Width, inspection/stimulus (inches)	Location, stimulus attended to
1	1	preferred hand
2	1	non-preferred hand
3	2.5	preferred hand
4	2.5	non-preferred hand

Apparatus. The apparatus consisted of a standard stimulus 1.5 in. wide (1.5 in. thick and 60 in. long), a comparison stimulus in the form of a wedge tapering from 2.88 in. to 1 in. at a rate of 1/32 in. per inch, and 2 pairs of inspection-stimuli, a 1-in. pair and a 2.5 in. pair. A moveable H-shaped rider, which kept *O*'s great finger and thumb directly opposite each other, rode on a wire which ran down the middle of the wedge. The top of the wedge was calibrated to permit readings to the nearest 1/32 in. The stimulus-objects (standard and comparison) were placed on tables 38 in. high.

Procedure. After the instructions were read to *O,* he was blindfolded and led between the standard- and the comparison-stimuli (the wedge). *E* placed the thumb and great finger of *O*'s non-preferred hand across the 1.5-in. width of the standard, and the thumb and great finger of the preferred hand across the width of the wedge. He was then allowed to explore the entire length of the wedge and the standard before making two practice-judgments. These judgments were made by holding the standard with one hand and finding a point on the wedge judged equal in width to the standard. Each judgment was begun with the hand placed alternately on the wedge 8 in. above or below the point of objective equality.

The practice-judgments were followed by four pre-inspectional judgments with the same procedure. To minimize the use of positional cues, the wedge was randomly moved back or forth within a distance of 12 in. between trials.

After the fourth pre-inspectional match, *O* removed both hands from the apparatus and was told that he would rub blocks of wood with both hands before making further judgments. He was then instructed to concentrate on impressions from one hand and this instruction was repeated for emphasis.

O's hands were then placed on the two inspectional stimuli and he moved his hands back and forth along the stimuli for 45 sec. Following the inspectional period, he made four post-inspectional judgments. The difference between the mean of the pre-and post-inspectional judgments reflects the effect of attention.

Results. The number of *O*s with judgments in the predicted direction, in the opposite direction, and showing no change is given in Table II for each of the groups.

TABLE II

DIRECTION OF KINESTHETIC AFTER-EFFECT FOLLOWING TWO-HAND EXPOSURE TO THE INSPECTION-STIMULUS WITH ATTENTION DIRECTED TO ONE HAND

Group	Predicted	Opposite	No change
1	7	5	0
2	6	6	0
3	8	4	0
4	7	4	1
Total	28	19	1

Though the results tend in the direction predicted by the attention-satiation hypothesis, they are not, however, as positive as those found by Köhler and Adams, where 35 of 42 *O*s showed effects in the predicted direction, while the remaining *O*s (7) showed no change. A Chi-square analysis of the present data based on all 48 *O*s and an even-split null hypothesis between predicted and opposite direction does not, however, allow for rejection of the null hypothesis at the 0.05 level of significance ($X^2 = 1.72$, $df = 1$, $p = 0.20$).

Experiment II. In view of the weak support for the attention-satiational hypothesis in the results of Experiment I, a second replication was attempted. The same general design was used, but the instructions were modified to increase the emphasis on attention, and in addition, *O* was periodically reminded to attend to the appropriate stimulus during the course of the inspectional period.

Method. The procedure was the same as that used in Experiment I except for a slightly larger sample (*N*-52 in four groups of 13 each) and the differences noted above.

Instructions. Instructions designed to emphasize attention were as follows:

> While you are rubbing the blocks, I wish you to pay close attention to the sensations from the fingertips of the right [left] hand as you move it back and forth. It is very important for you to concentrate on the fingertips of the right [left] hand. While rubbing with both hands, keep saying to yourself, "Pay close attention to the feelings in the right [left] hand."
>
> "To help you further in keeping your attention directed to the fingertips of your right [left] hand, I will say to you from time to time as you rub, "Pay close attention to your right [left] hand." [*E* said this every 10 sec.] Remember the important part of this experiment is your paying close attention to the sensations in the fingers of your right [left] hand."

Results. The results of Experiment II are shown in Table III. It can be seen that, with the exception of Group 3, the results again tend in the direction predicted by the attention-satiational hypothesis. The number of cases lying in the 'wrong' direction again makes for results less convincing than those reported by Köhler and Adams.

The marked preponderance of results in the opposite direction for Group 3 was unexpected on the basis of the hypothesis, the results of Köhler and Adams, and the results of Experiment I where this group showed the strongest results in the predicted direction. Because the discrepant results of Group 3 could not be explained, we decided to use another group of *O*s under the same conditions. The results of the repetition are shown in Table III as Group 3a. The results of this group are in line with the results of the other groups. On the basis of the evidence available, it appears plausible that the discrepant results of Group 3 represent chance deviation from results that might more frequently be obtained under these conditions.

TABLE III

DIRECTION OF KINESTHETIC AFTER-EFFECT FOLLOWING TWO-HAND EXPOSURE TO SATIATION WITH ATTENTION DIRECTED TO ONE HAND

(Experiment II)

Group	Predicted	Opposite	No change
1	8	4	1
2	8	4	1
3	3	10	0
(3a)*	(9)	(3)	(1)
4	10	3	0
Total (with 3)	29	21	2
(with 3a)	35	14	3
(with 3 and 3a)	38	24	3

* Replication of the conditions of Group 3 with a different sample.

A Chi-square analysis of the first four groups of *O*s does not allow for rejection of the null hypothesis at the 0.05 level ($X^2 = 1.23$, $df = 1$, $p = 0.30$). If the analysis is based on all the data available, *i.e.* Groups 1, 2, 3, 3a, and 4, then the null hypothesis can be rejected at the 0.10 level ($X^2 = 3.14$, $df = 1$). With the substitution of Group 3a for Group 3 the null hypothesis can be rejected at the 0.005 level ($X^2 = 8.01$, $df = 1$).

Since there was no significant difference between the observed proportions in Experiments I and II, all the available data (including Group 3a in Experiment II) were combined for a single significance test. The combined frequencies are 66 cases in the predicted direction, 43 cases in the opposite direction, and 4 cases showing no change. An analysis of these frequencies allows for rejection of the null hypothesis at the 0.05 level ($X^2 = 4.66$, $df = 1$).

The results of the two studies taken together tend to support the attention-satiational hypothesis, though the support is weaker than that obtained by Köhler and Adams. The difference between their experiment and the combined results of these experiments was tested for significance by a Chi-square test of differences in proportion of cases supporting the hypothesis. This difference was significant at the 0.001 level ($X^2 = 14.02$, $df = 1$). Despite the fact that both sets of studies support the hypothesis, there is a significant discrepancy between them in the strength of this support.

The reasons for the discrepancy between the studies is not readily apparent. Perhaps the use of a smooth wedge and finer calibration, contributed to the differences in some way. Differences in the wording of the instructions or samples used may also have been factors.[4]

Experiment III. In Experiment III the effect of attention is evaluated under conditions where only one hand is exposed to an inspectional stimulus.[5] In addition, a comparison was made between the moving and stationary exposure to the inspectional stimulus.

The extension of the satiational theory to include the effect of attention affords theoretical justification for expecting a difference between the moving and stationary conditions, since the moving condition is more likely to direct attention to the stimulus. Köhler and Dinnerstein report that stationary exposure can produce a *KAE,* but they point out that "we do not know whether the size of the results depends on such variations as to procedure."[6]

Design. The *O*s (150 undergraduate students) were divided into 6 groups of 25 each to allow for combinations of 3 sets of instructions (attention, distraction, and control) and 2 methods of exposure to the satiational stimulus (movement vs. stationary) as outlined in Table IV. Each group consisted of 8 women and 17 men.

Apparatus. The standard stimulus and the comparison-wedge were the same as those used in Experiments I and II. Only one inspectional stimulus, 0.5 in. wide, was used and it was placed on the side of *O*'s non-preferred hand.

Procedure. After making two practice-judgments, each *O* made four pre-inspectional judgments. Following this, the special instructions with respect to attention and stimulus-exposure were read to *O*. There followed a 60-sec. inspection after which *O* made four post-inspectional judgments. All his judgments were started with the hand at the extreme end of the wedge, with alternation between the wide and narrow ends on alternate trials. The difference between the means of the pre- and post-inspectional judgments was used as the measure of the *KAE*. An increase in apparent width of the standard after inspection constitutes a *KAE* under these conditions.

---

[4] The exact wording of the instructions was not given in Köhler and Adams' report.

[5] The use of a single inspectional-stimulus held in one hand is common practice in experiments on kinesthetic after-effects, *e.g.* J. P. Charles and C. P. Duncan, The distance gradient in kinesthetic figural after-effect, *J. exp. Psychol.,* 57, 1959, 164-170; Asenath Petrie, Walter Collins, and Philip Solomon, The tolerance for pain and sensory deprivation, this JOURNAL, 73, 1960, 80-90; H. H. Spitz and R. S. Lipman, Reliability and intercorrelation of individual differences on visual and kinesthetic figural after-effects, *Percept. mot. Skills,* 10, 1960, 159-166.

[6] Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinaesthesis, in *Miscell. Psychologica Albert Michotte,* 1947, 196-220, esp. 204.

Instructions. The instructions read to *O* prior to the inspectional period are presented below. The instructions read to the control groups were common to all the *O*s. For the experimental groups (attention and distraction), special instructions were inserted between the first and last paragraphs of the instructions. Instructions for the 'movement' and 'stationary' groups differed only as necessary to make them relevant.

TABLE IV

OUTLINE OF EXPERIMENTAL CONDITIONS FOR EXPERIMENT III

Group	Attention condition	Inspection-stimulus
1	attention	stationary
2	attention	move
3	control	stationary
4	control	move
5	distraction	stationary
6	distraction	move

TABLE V

MEAN KINESTHETIC AFTER-EFFECT IN INCHES

	Control	Attention	Distraction
Rub	.120	.075	.058
Hold	.058	.066	.061

Control with movement. "Now you are going to rub back and forth on a block of wood. I will tell you when to start and stop and I will show you how fast I want you to rub. You are not to talk during this period. [Specific instructions for 'Attention' and 'Distraction' groups, inserted here.] Immediately after I tell you to stop rubbing, you will make some more judgments of the width of a block of wood. Try to make each judgment within 15 sec."

Attention with movement. "Now this is important. You are to pay very close attention to the wood block while you are rubbing back and forth on it. While you are rubbing, I would like you to try to be constantly aware of the sensations of contact between your hand and the block of wood. I will help you to keep your attention on the block of wood by reminding you from time to time to pay attention to it." [Every 15 sec. *E* says the following to *O* "Pay close attention to the block."]

Distraction with movement. "While you are rubbing the block of wood I wish you to recite the alphabet backwards. Be as fast and as accurate as you can. Concentrate on speed and accuracy. While reciting the alphabet backwards you are to continue rubbing back and forth on the wooden block. Do not stop rubbing the block while you are reciting the alphabet."

Results. The mean *KAE* in inches is presented for each of the six groups in Table V.

The analysis of variance indicates that the over-all effects of attempts to manipulate attention were not statistically significant. The results offer, however, some support to the attention-satiational hypothesis in that the distraction resulted in the smallest *KAE,* and 'movement' produced a larger *KAE* than 'stationary.' It is of interest to note that even though the interaction between movement-stationary and attention was not statistically significant, the difference between the control-stationary and the control-movement groups was significant at the 0.05 level ($t = 2.88$).

A surprising result was the higher *KAE* for the control-movement than for the attention-movement group. Since the mean for the control-movement group was somewhat out of line with the other means and with the predictions made, another group was run under the control-movement condition as a check. The mean *KAE* for the second control-movement group was 0.135 in., which is close to the mean of 0.12 in. found for the original group. A *KAE* of about the same magnitude has also been reported by Charles and Duncan with a sample tested under similar conditions.[7] These independent findings suggest that the relatively large *KAE* found for the control-movement group in the original experiment may represent a real effect.

The second control-movement group allowed for another test of the difference between the control-stationary and the control-movement groups noted above. The difference between the original control-stationary group and the repeat control-movement group is significant at the 0.01 level.

Discussion. The results of Experiment III, together with independent corroborative evidence from other studies, suggest that the control-movement condition produces a larger *KAE* than the other conditions. It is suggested that attention may be optimal under this condition. In terms of the attention-satiational hypothesis it may appear paradoxical for the control-condition to yield a larger *KAE* than the attentional condition. If, however, the experimental situation is seen from the point of view of *O,* this paradox loses some of its force. A naïve *O* is told to rub his hand back and forth along a block of wood while blindfolded. In the absence of visual stimuli, he is exposed to novel stimulation involving motion of his hand. It seems very likely that under these conditions his attention will be strongly focused on inspection without special instructions to do so.

In the attentional condition of Experiment III, an attempt was made to direct attention to a stimulus which without instructions may already receive optimal attention. The special instructions may even have had a distracting effect, since *O* now had to attend to the process of attending and also to the voice of *E* telling him to attend during the inspectional period.

---

[7] Charles and Duncan, *op. cit.,* 166.

Though the attentional condition may have had a distracting effect, it still produced somewhat greater effects than the condition of distraction. The superiority of the control-movement condition over the control-stationary condition suggests that the 'hold-move' variable has an effect on the magnitude of the *KAE*. The 'hold' conditions did not show as much effect in the 'attention' and 'distraction' conditions. This may be due to the fact that the *KAE* under these conditions was already small and therefore subject to less decrement.

This line of thinking about Experiment III suggests an analysis of the effect of attention on the *KAE* found in the first two experiments reported above, and also in the Köhler and Adams experiment. Those experiments show a differential effect of satiation from the two hands as a function of direction of attention to the stimulus in one hand during the inspectional period. It is not clear from the results, however, just where the differential effect comes from. Does the attention to the stimulus in one hand result in more satiation from that hand than is characteristic for the satiation of one hand? Or does the distraction of attention from the stimulus in one hand, due to attention to the other, result in less satiation than is characeristic of the one-hand situation? Or is there an interaction of both effects?

The results of Experiment III suggest that the direction of attention to the hand (movement-condition) does not increase the effect of the inspectional stimulus over what it would be without the direction of attention; if anything, it tends to decrease the effect. It appears more likely that the differential effect is due to decrement in satiation from the hand which is *not* attended to.

A simple quantitative model may be helpful in spelling out the relationships between the one- and two-hand conditions of inspection. Assume that the amount of satiation, and consequently the size of after-effect, is related to the degree of attention directed to the inspectional stimulus. Assume further that one-handed inspection (movement) without instructions to attend to the inspectional stimulus produces optimal attention. Optimal attention is arbitrarily defined as X units of attention. Under the two-handed condition, and in the absence of instructions to attend to the stimulus in either hand, it is assumed that there is an even split of the available attention, so that each hand receives 0.5 X units of attention. Since, under this condition, there is no differential in attention between the stimuli in the two hands, no differential in *KAE* is expected. If instructions to attend to the stimulus in one hand are now introduced in the two-handed condition, a redistribution of attention should result such that the stimulus in the hand attended to now receives more (*e.g.* 0.7 X units) attention than the stimulus in the hand not attended to (*e.g.* 0.3 X units). It is this differential which produces the differential in *KAE* shown in the two-handed experiments. Note that it is possible for this differential to occur even when the attention directed to one of the hands is not the optimal attention, *i.e.* X units.

The total of evidence from the three experiments tends to support the attention-satiational hypothesis. An implication of this hypothesis is that a central factor, attention, which is distinct from the central effect of the inspectional stimulus itself, can influence the effect of the inspectional stimulus. This implication has a bearing on the interpretation of other studies of figural after-effects. For example, Wertheimer's findings of a relationship between metabolic efficiency and figural after-effects may possibly be mediated by differences in attention related to differences in metabolic efficiency.[8]

Though the exact mechanism of how attention exerts an influence on satiation is not clear, the results of recent work in neurophysiology suggest a possible mechanism. From the work of Hernandez-Peon, Scherrer, and Jouvet on the effect of attention on transmission of afferent impulses to the cortex, it appears that the physiological effect of a stimulus that is not attended to does not reach the highest cortical levels.[9] Given a fixed period of exposure to a stimulus, it would be expected that, when attention is high, the afferent impulses from that stimulus will reach the cortex a greater percentage of the time than when attention is low. If satiation takes place in the cerebral cortex, as is generally assumed, then the degree of satiation, and consequently the size of the *KAE,* would be diminished as fewer afferent impulses from the satiational stimulus reach the cortex. Thus, the role of attention appears as a selective filter determining the probability of the arrival of afferent impulses to the cortex for a given stimulus.[10]

SUMMARY

Three studies of kinesthetic after-effects suggested by the attention-satiational hypothesis of Köhler and Adams were reported. It was predicted that if two inspectional stimuli of equal width are held, one in each hand, with attention directed only to one of the stimuli, then a greater after-effect would result from the stimulus attended to. The results of two experiments supported the prediction. Though each of the experiments fell short of statistical significance, an analysis based on the combined data showed a statistically significant effect of attention.

---

[8] Michael Wertheimer, Figural after-effects as a measure of metabolic efficiency, *J. Personal.*, 24, 1955, 56-73.

[9] Raul Hernandez-Peon, Harold Scherrer, and Michael Jouvet, Modification of electric activity in cochlear nucleus during "attention" in unanaesthetized cats, *Science,* 123, 1956, 331-332.

[10] A filter theory of attention has been elaborated by D. E. Broadbent (A mechanical model for human attention and immediate memory, *Psychol. Rev.* 1957, 64, 205-215); *Perception and Communication,* 1958, 190-196, 251-256, 297-301.

A third experiment involved manipulation of the attentional variable in a situation where only one hand was exposed to an inspectional stimulus. An attempt was made to manipulate attention for two groups by (a) instructing *O* to attend closely to the inspectional stimulus, and (b) by having *O* engage in a distracting task to detract attention from the stimulus. For a third group (control), no special effort was made to control attention. Each of these groups was further subdivided into a 'stationary' group and a 'movement' group according to whether *O* held the stimulus or moved his fingers back and forth along it. This variation allowed for a test of the hypothesis that inspecting a stimulus during movement would produce greater after-effects, since it constitutes a form of stimulation more demanding of attention. Results of an over-all analysis of variance revealed no significant differences between the groups, though the hypotheses were generally supported by the direction of the results. The nature of the relationship between attention and the after-effects is discussed.

NAVAHO WORD-ASSOCIATIONS

By SUSAN M. ERVIN, University of California, Berkeley,
and HERBERT LANDAR, Los Angeles State College

It has been found repeatedly with educated adults who speak Indo-European languages that most responses in the word-associations are in the same grammatical class as the stimulus-words.[1] Children's associations, on the other hand, are characteristically sequential. The major shift, for both American and French children, comes in the first few years of school.[2]

The current study of Navaho associations is principally designed to examine the role of grammatical classes in association. First, the levels of paradigmatic responses, *i.e.* responses in the same grammatical class, will be compared with those obtained in other studies. Secondly, internal comparisons will be made, to see whether classes occurring in final position in sentences have higher frequencies of paradigmatic responses. Thirdly, the relation between contextual variety and sequential responses will be tested with a Navaho system of classes with highly restricted contexts. Fourthly, the level of commonality will be compared with the levels found in other studies. The physical isolation of Navaho residences from each other, and lack of mass-media should decrease commonality. Finally, individual differences in schooling and Navaho dominance in bilingual *S*s will be related to associative behavior.

METHOD

Subjects. Nine cases were tested near Shiprock, New Mexico, and 28 near Fort Defiance, Arizona. These areas are within the Navaho reservation, and are sur-

---

* Received for publication July 27, 1961. This research was sponsored by the Southwest Project in Comparative Psycholinguistics of the Social Science Research Council. The second author did linguistic analysis while employed by the Department of Public Health, Cornell University Medical College. We are grateful to the Navaho staff of the Navaho-Cornell Clinic, and particularly to William Morgan, Assistant in Anthropology of the Cornell-Navaho Field Health Research Project, for services as linguistic informants, and to Professor Harry Hoijer of the University of California for linguistic advice.

[1] The most detailed report was that of Arthur Wreschner, Die reproduktion und assoziation von vorstellungen, *Z. Psychol. Ergb.*, 3, 1907, 329-599.

[2] R. W. Brown and Jean Berko, Word-association and the acquisition of grammar, *Child Develop.*, 1960, 31, 8; S. M. Ervin, Changes with age in the verbal determinants of word-association, this JOURNAL, 74, 1961, 361-372; M. R. Rosenzweig and and R. Menahem. Age, sexe, et níveau d'instruction comme facteurs déterminants dans les associations de mots, *Année psychol.*, 1962, 62, 45-61.

rounded by isolated clusters of hogans where Navahos live in traditional style, herd sheep, and congregate for rodeos, 'pow-wows,' and ritual purposes. Most *S*s were tested indoors in a room with modern furnishings. At least 8 of those tested were clearly monolinguals, 16 knew some English, and 13 were English-dominant bilinguals, judging from their relative speed in naming 120 pictured objects in the two languages. Ages ranged from 17 to 70 yr. Twenty-five were women.

Instructions. The instructions were tape-recorded in Navaho, but an interpreter was present in case of misunderstanding.

> Here is a sheet of paper with some words written on it. The instructor will call out a word for you and he wants you to repeat the word and add another word to it. The way you add another word is something like this. The first word is 'cat.' Think of a cat, and think that he drinks milk, and that he kills birds and mice, and he likes to be near a fire, and he has whiskers. You can find a lot of words connected with a cat. This is just an example. Find words connected with the words that I say. Don't hold back, go right ahead. I'm going to give three more words for you now. 'Wagon': with a wagon the first thing I think of is 'horse.' The next word is 'cloud': then I think of 'rain.' For 'money' I think of 'food.' You can say any word, even though it's not connected with the thing; you can say any word that comes to mind. As soon as you repeat the word, say another word.

During testing, the *E*s called out the words, heard them repeated to guarantee comprehension, and tape-recorded the responses. When the responses were too long, S was urged to shorten them, and more examples were supplied.

Materials. One hundred and fourteen items were presented, drawn in scrambled order from different grammatical classes. To clarify later discussion, a brief description of Navaho grammatical classes is necessary.[3]

Fifty stimulus-words were nouns or nominal compounds. These words or compounds are defined partly by syntactic function and partly by inflection. Nouns are single words which are inflected for possession but not for number. Compounds and one-word nouns usually lack inflectional suffixes to provide a basis for sound-similarity, but 11 items referred to parts of the body, or relationships ('wife,' 'father') which require a possessive prefix (*bi-má* = 'his mother,' *qa-má* = 'somebody's mother').[4] Such items tend to elicit responses in which either the base or prefix is repeated.

Nominal compounds have a diverse internal composition, and may be made up of various sequences of parts of speech, but they all occupy the position in sentences of ordinary simple one-word nouns. Nominal compounds may be formed of two nouns ('gas' = *chidí bi-tooq* = 'car its-water'), a noun with an enclitic ('grandmother' = *qamá sání* = 'somebody's-mother old'), or a noun with a verb ('sugar' = *qásbĩĩh Ləkan* = 'salt it-is-sweet').

One word was a numeral, *qashdlaq* ('five'), a member of a nominal subclass defined by syntactic function.

---

[3] Further information about Navaho is available in Harry Hoijer, The Apachean verb, Part I: Verb structure prefixes, *International J. Am. Linguistics,* 11, 1945, 193-203; Part II: The prefixes for mode and tense, *op. cit.,* 12, 1946, 1-13; Part III: The classifiers, *op. cit.,* 51-59; Part IV: Major form classes, *op. cit.,* 14, 1948, 247-259; Part V: The theme and prefix complex, *op. cit.,* 15, 1949, 12-22; Robert W. Young, *The Navajo Yearbook,* 1958, 197-228.

[4] The following conventions have been used for transcribing Navaho sounds 'L' refers to a voiceless 'l', a lateral spirant; an acute accent refers to raised pitch; double vowels are of longer duration; 'ə' refers to schwa, the vowel in 'gum'; 'q' refers to a glottal catch; and ~ above vowels refers to nasalization. All underlined words are literal stimuli or responses, whereas words in quotation marks are approximations or translations.

Thirty-four stimulus-items were simple verbs. Verbs may be transitive or intransitive as in English. Translation of some intransitive verbs, however, reveals that adjectival meaning is involved. There is in fact no separate formal category of adjectives in Navaho, as there is in most Indo-European languages. The 50 simple verbs consisted of a prefix and stem. In addition, 4 verbs were more complex, including additional prefixes. The typical Navaho verb consists of a long string of prefixes plus a stem.

Six verbs exhibit a special property which has attracted attention in research on language and cognition.[5] To say 'hand it to me' one must select among eleven alternative stems. The selection is based on the physical properties of the thing you want. Many verbs of handling have a set of alternative stems. The six verbs used as stimuli were the six different ways of saying "it's sitting there," used respectively when one speaks of round bulky objects, fluffy things, flat flexible objects, things in containers, long rigid objects, and long flexible objects. As far as we could ascertain, this system is entirely consistent semantically, so that a child can use it correctly for new objects. Metaphorical extensions like "he brought the news," employ the round bulky object-stem. Necessarily these stems occur in comparatively restricted environments, since a noun that can occur with one will not usually occur with another. The exception is the name of an object that can change form: a rug unrolled is a flat flexible object but rolled up is a long object. Hereafter these 6 verbs will be called specialized verbs.

One stimulus-word was a particle, *nishqnáájí* ('right'), and four were relational expressions. These can be translated as 'high,' 'fast,' 'down,' and 'carefully' or 'slowly.' The last three end with an enclitic particle, *-go*. These terms tend to be used as modifiers before verbs.

A sentence may consist of merely one word, a verb (*e.g. Lichíiq,* 'it is red,' or *jiní,* 'they say it') which may be translated with the subject included, or with both the subject and object included. A basic pattern of longer utterances, which contain nouns as well as the verb-stem and prefixes, is subject, object, and verb, in that order (*e.g. gaLbáhí mãqii dayiiLtsã,* 'rabbit coyote they-saw-him').

Analysis. Two procedures were used for initial classification of responses. A Navaho informant identified which responses readily suggested a context in which they could substitute for the stimulus-word, and whether they could precede or follow the stimulus-word in utterances. Also, each word was classified as to grammatical class. In the subsequent analysis, a response is considered paradigmatic if (1) the item is in the same grammatical class as the stimulus-word, and (2) it does not contain the stimulus-word. An exception to the second rule was made in certain cases. Some very common compounds such as 'sugar' (*qáshĩĩh Ləkan* = 'salt it-is-sweet') and 'puppy' (*Léécháāqi yázhí* = 'dog young') contain their stimulus-words. It was decided that such compounds may be psychologically paradigmatic if the Navahos treated them as such.[6] Individuals who gave predominantly paradigmatic responses on unambiguous items, and others who gave predominantly syntagmatic responses were separated. These compounds were considered paradigmatic if the first group gave them as responses but none of the second group did. If some of each gave them, they were considered ambiguous and omitted from the computations.

---

[5] Howard Maclay, An experimental study of language and non-linguistic behavior, *Southwestern J. Anthrop.,* 14, 1958, 220-229; J. B. Carroll and J. B. Casagrande, The function of language classifications in behavior, *Readings in Social Psychology,* 1958, 18-31.

[6] Similar constructions in English, such as 'table'–'tablespoon,' or 'sweet'–'bittersweet' were omitted entirely in Ervin, *loc. cit.,* but were judged paradigmatic by Brown and Berko, *loc. cit.* and Rosenzweig, *loc. cit.*

Responses were considered sentences if they were quite long or included more than one basic constituent in addition to the stimulus-word. Except where indicated, sentence-responses were excluded. Both stimulus- and response-words which could be in more than one grammatical class were excluded. The paradigmatic score for each stimulus-word consists of the percentage of remaining responses which are paradigmatic. Exclusion of ambiguous or complex stimuli yielded 100 analyzable stimuli.

RESULTS

Primary responses. The level of commonality for primary responses was considerably lower than the level obtained on college students.[7] If 47 items which correspond in the norms of Russell and Jenkins and in the Navaho data are compared, the median commonality for the Navaho items is 17% and for the English items is 45%. The low level is not a result of the bias arising from selection of corresponding items from the Kent-Rosanoff list, since the median for the total list of 100 words given to the Navaho was 17%.

A very large proportion of the primaries were paradigmatic. Of the 47 corresponding words, 5 in English and 4 in Navaho produced syntagmatic primaries, 2 in English and 5 in Navaho were ambiguous, and the rest were paradigmatic.

It is of interest to compare the actual primaries of the translated terms, though there are, of course, artifacts in such comparisons. Twenty-four of the 47 items have corresponding primaries, when ties are included if either corresponds. Correspondences seem to be concentrated in certain types of words, chiefly terms which have natural opposites, such as terms with adjectival meaning, nouns with a contrasting mate (man–woman, boy–girl, hand–foot), and culturally corresponding pairs of tools (needle–thread, hammer–nail). Discrepancies appear if a term has several meanings in one language (foot and shoe are the same in Navaho), there are compounds containing one of the terms (sugar = sweet salt, and chili = bitter herb), or there are cultural differences in the use of the referent. The latter of course is the principal case here, and probably accounts for the fact that correspondences were found to be less for Navaho and European languages than with the latter languages, as reported by Rosenzweig.[8] Some examples of culturally based differences seem to be these: English *bread–butter* vs. Navaho 'bread'–'fried bread', English *blossom–flower* vs.

---

[7] W. A. Russell and J. J. Jenkins, The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test, Studies in the role of language in behavior, Univer. Minn., Techn. Report No. 11, N8 onr 66216, 1954.

[8] Rosenzweig, Comparisons among word-association responses in English, French, German, and Italian, this JOURNAL, 74, 1961, 347-360.

Navaho 'blossom'–'pollen' (ritually important), English *house–home* vs. Navaho 'house'–'hogan,' English *whiskey–drink* vs. Navaho 'whiskey'–'it's no good,' (prohibited by the tribal council on the reservation).

If we turn to the total list of 100 words, we find nearly all the primaries to be paradigmatic. The exceptions are 'corn'–'white,' 'whiskey'–'it's no good,' 'noisy'–'children,' 'ripe'–'corn,' 'bitter'–'chili' (bitter–herb), 'angry'–'cattle,' and the six specialized verbs. These verbs, with their great restriction in permitted contexts, all yielded syntagmatic primaries consisting of the nouns which most typically precede each of them: 'stone,' 'wool,' 'blanket,' 'water,' 'log,' and 'rope.' In addition to the opposites and tools mentioned above from the Kent-Rosanoff list, the words eliciting paradiagmatic primaries with the highest commonalities were terms of kinship, parts of the body, natural pairs ('lightning'–'rain,' 'moon'–'sun'), and 'five.'

Paradigmatic responses. The proportion of paradigmatic responses varied with class. The largest proportion occurred among nouns, with a mean of 59% paradigmatic responses. Verbs had a mean of 52% paradigmatic responses, modifiers 42%, and specialized verbs 29%. To the word 'five' 47% paradigmatic responses were given, if only numerals are considered paradigmatic. These levels of paradigmatic response may be compared to other findings. Brown and Berko reported means of 30%, 53%, 55%, and 81% paradigmatic responses for first-, second-, and third-grade children and for adults respectively, when mass-nouns are omitted.[9] Ervin, with a list containing fewer adverbs, obtained 37% for kindergarten and first-grade children, 62% for third-grade children, and 70% for sixth-grade children.[10]

The principal syntagmatic responses to nouns and modifiers were words which usually follow them in utterances. The principal syntagmatic responses to verbs were nouns which could precede the verbs. Since Navaho permits adjacent occurrence of nouns in sentences, some items classified as paradigmatic could occur contiguously in discourse. In the second English study mentioned above, responses in the same class which could be syntagmatic were omitted (*e.g. front–door*). The prevalence of this pattern in Navaho made such omissions impossible.

Class-analysis. It is conceivable that the levels of paradigmatic response obtained above could be derived from random responses to stimuli. In Table I it can be seen that this is not the case. Nouns and specialized verbs tend to elicit nouns. Ordinary verbs elicit verbs, and modifiers elicit both other modifiers (chiefly opposites) and verbs.[11]

---

[9] Brown and Berko, *op. cit.,* 8.

[10] Ervin, *op. cit.,* 367.

[11] Discrepancies from the paradigmatic percentages reflect the different base used in the two calculations.

Individual differences. There were very extreme differences in preferences for paradigmatic responses. Twenty-six individuals for whom data on schooling was available were chosen, and a sample of 55 responses was taken for each person. It was found that 3 persons had fewer than 6% paradigmatic and 3 had more than 94% paradigmatic responses. There was a small and unreliable tendency for those who had schooling or were English-dominant to give more paradigmatic responses.

Sentence-responses appeared in 9 of the 11 persons who were Navaho-dominant or monolingual and had no schooling, but only 3 of the remaining 15 people gave such responses. Thus the concept of the 'word' as a short unit seems to be acquired primarily though reading, even if it is reading in another language of quite different structure. Since the separation of the sentence-responses from the syntagmatic responses was somewhat arbitrary, there is some justification for pooling them. If this is done, the uneducated, Navaho-dominant group gives significantly fewer paradigmatic responses at the 2.5% level by the Mann-Whitney test.

TABLE I

CONDITIONAL PROBABILITY OF CLASSES IN ASSOCIATION

Response-class	Stimulus-class noun	specialized verb	verb	modifier
Noun	.53	.51	.33	.11
Verb	.37	.38	.50	.43
Modifier	—	.01	.07	.39
Other	.02	.02	.01	.02
Sentence	.08	.08	.08	.06

DISCUSSION

Paradigmatic responses. The level of paradigmatic responses is considerably lower for the Navaho than for educated adults in other cultures. The explanation is not obvious. Since a similar difference has been found in France between educated and uneducated samples,[12] it is probable that the educational process accounts for these differences among groups and for the changes with age.

But why might this be? One difference appears in this study as well as in the research on changes with age—education reduces the length of responses. Both sentence-responses and responses which include the stimulus-word are reduced. Paradigmatic responses are thus more possible because sentence-responses or phrase-responses containing the stimulus-word usually cannot be paradigmatic, or at least are not so judged in the analysis. But the differences remain when sentence-responses are omitted.

---

[12] Rosenzweig, *Année psychol., op. cit.*

An explanation of group-differences must be based on a more general account of the development of paradigmatic preferences. It is clear that paradigmatic responses reflect possibilities of substitution in specific contexts. When English-speaking *Ss* have been asked to produce sentences and make substitutions for particular words (also used as stimulus-words for association), the substitutions they make correspond overwhelmingly to paradigmatic primaries. This is especially true of antonyms and other instances of high commonality.[13]

It may be expected that a prerequisite to the development of paradigmatic responses is sufficient contextual variety to reduce syntagmatic competition. The specialized verbs in Navaho provide a neat, grammatically defined test of this prediction. Since the 11 alternatives are contextually selective, the range of contextual variety is much less than that of ordinary verbs. The fact that all of these verbs produced primarily syntagmatic responses, and that these tended to be the nouns which usually precede the verbs, shows the importance of contextual variety.

If the difference between groups is based on an increase in size of vocabulary and hence in contextual variety with age or education, there should be more syntagmatic primaries offered by the children, the Navahos, and the uneducated Frenchmen. In fact, this was true only of the French sample. We have assumed that commonality in a group reflects frequency for an individual. This is not necessarily the case. For children, and for the geographically dispersed Navahos, idiosyncratic concentrations on certain sequences may result in syntagmatic responses that are not reflected in primaries. In the case of the specialized verbs, the syntagmatic primaries reflected the fact that because of grammatical regularities for all Navaho speakers the same contextual selectivity exists. Thus we cannot rule out differences in the degree of contextual variety as an explanation for differences among groups.

Given contextual variety, why do paradigmatic responses occur? Several explanations have been offered, which are not necessarily mutually exclusive: (1) Paradigmatic responses may become associated with the stimulus through contiguity. Contiguity, while it would not occur in normal overt speech, might appear when a listener anticipates what another person says. Words which normally occur in the same semantic and linguistic contexts would most often be erroneously anticipated, *e.g.* 'is the mattress . . . hard? soft?' This explanation alone could not account for the differences among groups. (2) Even though paradigmatic responses may not be the strongest associates, the task may be interpreted as one of returning like for like. Sometimes *S*s are overt about self-instruction ("should I give opposites?") and educational experience may give some practice in similar tasks. The very marked individual differences in paradigmatic preferences suggest directing and monitoring of his own pattern of response by some *S*s. Even third-grade children ask whether there is a right answer; they may thus hunt for some rule to follow. (3) As language is learned, concept-formation may occur regarding grammatical classes. This explanation may be appropriate in accounting for changes with age, in conjunction with the second point, but it seems least satisfactory in accounting for differences among adults.[14]

---

[13] Ervin, unpublished data.

There has been some suggestion of directional preferences in syntagmatic responses in English. It now appears that this preference may reflect the structural properties of English. In most English sentences the antecedent to a verb is a pronoun in the subject case. Since subjective pronouns do not usually occur in single-word utterances in English, it is much more likely that a syntagmatic response to a verb will be a word which follows the verb rather than one which precedes it. It has been argued that the reason nouns show paradigmatic dominance earlier than transitive verbs is that they occur more often in final position in sentences and hence have occurred less often with a sequel. Yet it can also be argued that nouns are associated with more competing syntagmatic words, since they may follow verbs, and both follow and precede adjectives (in copular sentences). This competition may leave a paradigmatic response as the strongest possibility. A good example is the word *coffee.* It elicits many contexts: *drink* (preceding), and *good, hot,* and *black* (following). *Tea* is the primary association. *Game,* on the other hand, has only one frequent contextual word—*play,* and this is its most frequent associate although it normally precedes.[15] Brown and Berko's evidence that adverbs are slow in development of paradigmatic dominance, though they typically occur in final position, supports the argument that the basic principle is contextual variety rather than order.[16]

Now the Navaho data buttress this alternative. Navaho verbs strongly elicited the nouns which precede them. The particularly high frequency of syntagmatic responses in the case of the specialized verbs points to contextual variety as a fundamental variable in association.

---

[14] The first explanation is discussed in Ervin, *op. cit.,* and the third by Brown and Berko, *op. cit.,* 1-4.

[15] Ervin, unpublished data.

[16] Brown and Berko, *op. cit.,* 8.

SUMMARY

The grammatical properties of Navaho word-associations have been examined for 100 stimulus-words. It has been shown that (1) commonalities were very low; (2) primaries tended to be paradigmatic except when contextual variety was limited; (3) with primaries, preferences for contrasts appeared as in English; (4) the proportion of paradigmatic responses out of total responses for each word was as low as among primary-school children; (5) the grammatical class of the response depended upon the class of the stimulus-word; (6) if a word had a restricted distribution, syntagmatic frequencies were especially high; (7) forward associations appeared to be no stronger than backward associations.

EFFECTS OF DIRECTION UPON SPEED OF MIRROR-DRAWING

By GEORGE H. ZIMNY, Marquette University

The present experiment was designed to determine the relative difficulty, as measured by time, of the vertical, horizontal, and diagonal directions involved in mirror-drawing.[1]

Method: Subjects. Forty-eight right-handed girls served as the *Ss*. They were randomly selected from a group of 303 right-handed girls in a third semester course in English, required of all Liberal Arts students.

Apparatus and materials. Whipple's mirror-drawing apparatus,[2] two stop-watches, an accumulative timer, and a stimulus-sheet (Fig. 1) were used in this study. The stimulus-sheet shows eight pairs of parallel lines (radii), each beginning ⅜ in. from their common center. The parallel lines of each radius were 3⅛ in. long and 1/16 in. apart and were divided at ⅛-in. intervals into 25 cells or parts. The radii proceeded in the eight cardinal directions. Fig. 1 shows the directions as found on the stimulus-sheet; *not* as reflected in the mirror.

Procedure. S was seated at a table facing the apparatus with the stimulus-sheet flat upon the table before her (with the result that the vertical direction required movements in toward and out from the body and the horizontal direction required movements to the left and right of the center of the body). *E* was seated to the left beside her with the timing devices out of *S*'s sight. After being shown the stimulus-sheet in the mirror, *S* was instructed to draw a line with a pencil between a specific set of parallel lines as rapidly and as accurately as possible. Beginning with the division or cell closest to the center, the pencil mark was required to pass through every one of them. If the pencil left the confines of the parallel lines, *S* had to reënter at the point of departure.

S was told which radius she was to trace and how it appeared in the mirror. *S* closed her eyes and *E* placed the point of the pencil at the beginning point—the division nearest the center of the sheet—and at 'Go,' she opened her eyes and began tracing the path. Each *S* traced the eight radii once, and the sequence of radii was counterbalanced among the 48 *Ss*.

Five measures were obtained from every *S* for every position: the time required to trace (1) the first half of every radius; (2) the second half; (3) the entire radius; (4) the time outside the boundary lines; and (5 the time within them. Time was measured in units of ⅕ sec.

---

* Received for publication September 29, 1960.

[1] N. V. Scheidemann, A five-pointed star-pattern for mirror-drawing, this JOURNAL, 63, 1950, 441-445; E. M. Siipola, Studies in mirror drawing, *Psychol. Monogr.*, 46, 1935, (No. 210), 66-77; G. M. Whipple, *Manual of Mental and Physical Tests*, 1910, 343-350.

[2] Whipple, *op. cit.*, 344f.

Results. The median of each of the measures for the eight directions is given in Table I.

(1) Inter-directional effects. To determine the effect of the three directions a Friedman Chi-square test was carried out for each of the five measures of response.[3] All the values were significant beyond the 1/10% level.

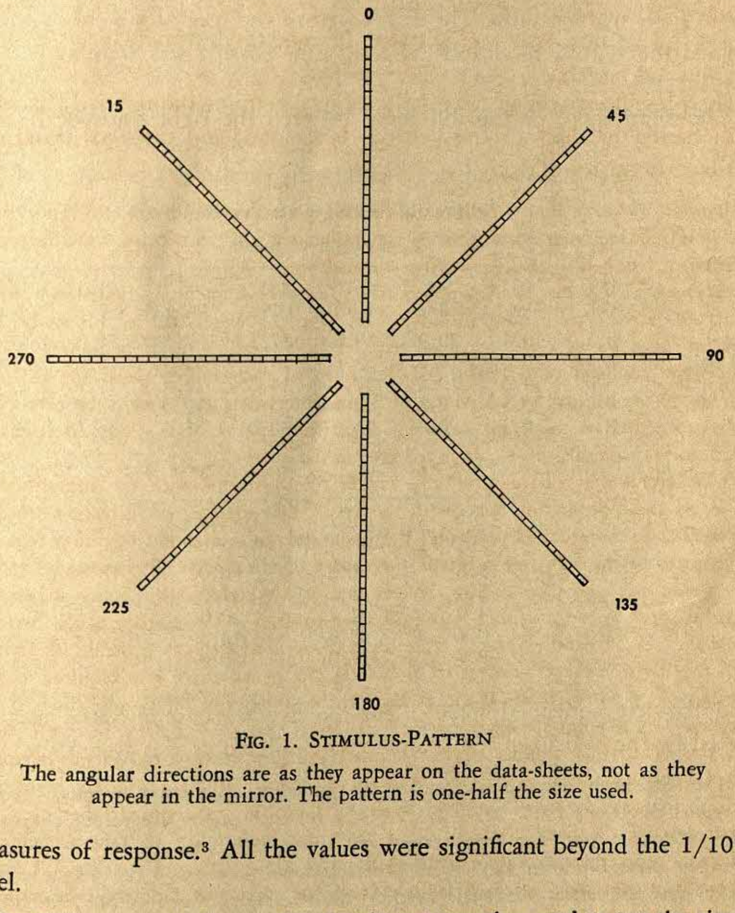


FIG. 1. STIMULUS-PATTERN

The angular directions are as they appear on the data-sheets, not as they appear in the mirror. The pattern is one-half the size used.

The Wilcoxon matched-pairs signed-ranks test was then used to test the significance of the differences between the vertical and horizontal and between the horizontal and diagonal directions for each of the five measures.[4] All the *z*s were significant beyond the ⅒% level. It was assumed, therefore, that the largest differences, those between the vertical and diagonal directions, were likewise significant beyond that level. For all measures of the response, the order of directions from least to most time required is: vertical; horizontal; and diagonal.

---

[3] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, 166-172.
[4] Siegel, *op. cit.*, 75-83.

(2) Intra-directional effects. The Wilcoxon test was carried out for the difference between the up vertical (0°), which is drawn out from the body, and the down vertical (180°), which is drawn in toward the body, for each of the five measures. The *z*s for four of the measures were significant beyond the 1/10% level and the *z* for time outside was significant at the 1% level.

The significance of the difference between the right horizontal (90°) and the left horizontal (270°) for the five response-measures was also tested by the Wilcoxon test. None of the differences was significant.

TABLE I

MEDIANS OF DRAWING-TIME (IN SEC.)

Time of:	Angular direction							
	180° Ver.	0° Ver.	90° Hor.	270° Hor.	135° Diag.	45° Diag.	315° Diag.	225° Diag.
Total	8.5	10.4	20.9	19.3	37.1	67.5	72.5	33.3
1st half	4.9	6.0	10.5	8.6	19.4	42.0	46.7	15.0
2nd half	3.6	4.5	10.5	10.5	14.7	22.9	21.1	12.0
'out'	0.3	0.5	2.4	2.9	5.0	10.0	13.0	3.7
'in'	8.4	10.2	20.0	16.6	30.0	55.0	56.7	31.1

Average		
both verticals	both horizontals	the four diagonals
9.5	20.5	60.0
5.4	8.5	37.0
4.0	11.2	20.0
0.1	3.0	14.8
9.5	17.0	50.0

The Friedman analysis demonstrated significant differences among the four diagonal directions (45°, 135°, 225°, 315°) for each of the five response-measures. The Wilcoxon test was then carried out using each measure for the difference between the right-down (135°) and the left-down diagonals (225°); the right-up (45°) and the left-up diagonals (315°); the two right diagonals (45° and 135°); and the two left diagonals (225° and 315°). Of the 15 *z*s obtained, only one was significant beyond the 5% level and this was discounted as being an artifact produced by the particular series of directions drawn by three high-scoring *S*s. The differences between the two up diagonals (45° and 315°) and the two down diagonals (135° and 225°) were significant beyond the 1/10% level for four of the measures and at the 1% level for the time 'outside' measure. Thus, it appears that the primary, if not sole, source of the significant difference found for the four diagonals lies in the up-down factor contained in the diagonals.

These analyses indicate the importance of the up-down (vertical) factor and the lack of importance of the right-left (horizontal) factor in mirror-drawing.

(3) First versus the second half. The Wilcoxon test was made for the difference between the time required to draw the first and the second halves of each of the eight radii and of the vertical, horizontal, and diagonal directions. The *z*s obtained for the vertical, up vertical, down vertical, diagonal, and each of the four diagonals were significant beyond the 1/10% level. In every case, the time to draw the first half was longer than the time to draw the second half.

The *z*s obtained for the horizontal and for the right horizontal were not significant (30% level); for the left horizontal the *z* was significant at the 1% level, but the time to draw the second half was longer than the time to draw the first half. In agreement with the findings for the intra-directional effects, it is the horizontal directions (right-left factor) that fail, in general, to show significant differences.

(4) Time 'in' versus time 'out' of the figure. The time within the boundaries of the figure was greater than the time outside for each one of the eight radii, and for the vertical, horizontal, and diagonal directions. Use of the Wilcoxon test showed that the difference between time 'in' and 'out' was significant beyond the 1/10% level in every case.

(5) Practice-effects. To investigate the effects of practice, two approaches were followed: (a) The *S*s traced the two vertical radii (0°and 180°) and the two horizontal radii (90° and 270°) in counterbalanced order. In the first approach the time spent on the first versus the second of the two vertical radii and the time spent on the first versus the second of the two horizontal radii were compared. (b) In the second approach, the eight trials (radii) given, in counterbalanced order, are grouped into the first, second, third and fourth blocks of two trials each and the time for each earlier block of two trials was compared with the time for each of the later blocks of two trials.

(a) For the first approach, Table II contains the medians for the five measures for the first and the second radii for the vertical and for the horizontal directions. None of the Wilcoxon *z*s was significant for the difference between the time spent on the first and on the second of the two vertical radii, thus indicating an absence of the effects of practice for this direction. For the horizontal radii, three of the *z*s were significant beyond the 1/10% level while the *z* for the time of the second half was significant at the 3% level and for 'time-inside' at the 4% level. The null hypothesis was rejected in every case. Since the time for the first radius was longer than for the second for every measurement, practice-effects were regarded as being present for this direction.

(b) For the second approach, Table III contains the medians for the five response-measures for the times (mean of two trials) for the first (Trials 1 and 2), second (Trials 3 and 4), third (Trials 5 and 6), and fourth (Trials 7 and 8) blocks of two trials. The Wilcoxon test was carried out for the difference between the time required for every pair of blocks of two trials for each of the five response measures.

For total time, the difference for each of the six pairs of blocks of two trials was significant beyond the 5/10% level. Since the total time decreased from the first to the fourth block, practice-effects are clearly demonstrated.

For time first half, the difference between the first and third block was significant at the 1/10% level, between the first and fourth blocks at the 3/10% level, and between the second and fourth blocks at the 2/10% level. The other three differences were not significant. The significant differences found between all non-adjacent blocks of trials coupled with the decrease in time indicates the presence of practice-effects in drawing of the first half of the figure.

For time second half, only two differences were significant. The difference between the second and third blocks was significant at the 3.5% level and between the second and fourth blocks at the 5/10% level. The four distributions are, however, very similar except for three extremely high scores in the second block. Consequently, it was decided to reject the two differences as being significant and to accept the null hypothesis. Thus, no practice-effects were found in drawing the second half of the figure.

TABLE II

MEDIANS OF DRAWING-TIME (IN SEC.) FOR THE FIRST AND SECOND RADII OF THE VERTICAL AND HORIZONTAL DIRECTION

Time of:	Direction			
	vertical		horizontal	
	Trial 1	Trial 2	Trial 1	Trial 2
total	9.6	9.0	23.0	18.8
1st half	5.5	5.9	12.6	8.3
2nd half	4.2	3.8	11.3	9.0
'out'	0.2	0.1	4.5	1.9
'in'	9.6	9.0	18.0	16.8

TABLE III

MEDIANS OF DRAWING-TIME (IN SEC.) FOR FOUR SUCCESSIVE BLOCKS OF TWO TRIALS EACH

Time of:	Blocks			
	first	second	third	fourth
total	32.5	38.8	27.2	19.1
1st half	18.8	20.0	13.0	9.8
2nd half	12.0	14.8	12.0	9.0
'out'	6.4	8.0	3.2	1.7
'in'	26.2	30.0	22.0	18.1

For time 'out,' the difference between the first and second and between the third and fourth blocks of trials was significant at the 5% level. The remaining four differences were significant beyond the $\frac{9}{10}\%$ level. With time decreasing from the first to the fourth block, the effects of practice were clearly present.

For time 'in,' the difference between the first and fourth blocks was significant at the 3% level and between the second and fourth blocks at the 1% level. Less time was spent at the fourth than at either of the other blocks, so evidence of some effects of practice was found.

Discussion: Direction. Our results clearly indicate that direction is an influential variable. Of the three directions, the vertical, in confirmation of Siipola, was traced most quickly;[5] and, in confirmation of both Siipola and Scheidemann, the diagonal direction was the most difficult.[6] Indeed, our results stand in close agreement with Scheidemann's statement that the diagonal is the only direction that tests mirror-orientation. An important implication of this: experiments employing mirror-drawing should take account of the difference in difficulty among the three directions as the results obtained will vary depending upon direction traced.

Perception of the figure in the mirror involved a reversal within the figure in the vertical direction but no reversal within the figure in the horizontal direction.[7] It was thus expected that the *S* would simply move in accordance with what she saw in the mirror while tracing the horizontal direction, thus spending less time inside than in the case of the vertical direction where a reversal did exist. It was found, however, that the vertical direction was the easier of the two for time inside, which means that the *S*s did better in the reversed than in the non-reversed situation. This finding is understandable when the greater over-all ease of the vertical in comparison to the horizontal direction is taken into account. It should be noted, though, that of the total time spend drawing each direction, a greater proportion of time was spent inside the vertical than inside the horizontal figure, indicating that the reversal produced relatively greater difficulty in drawing than did the non-reversal, as expected.

In this mirror-drawing task, an error occurs when the pencil leaves the boundary of a figure. As suggested by Siipola, *correction* of errors in the horizontal and diagonal directions is complicated by a reversal, due to viewing the figure in the mirror, whereas no such complication exists in correcting errors in the vertical direction.[8] It was predicted that more time would be spent drawing inside the vertical figure, where a reversal exists, than correcting errors outside the figure, where no reversal exists. This prediction was confirmed. It was also predicted that more time would be spent outside the horizontal figure correcting errors, where a reversal exists, than would be spent drawing inside the figure, where no reversal exists. This prediction was not confirmed, since the difference between time in and out, while significant, was in favor of time in. For the diagonal direction, where a reversal exists when drawing both inside and outside the figure, time in was significantly greater than time out. While the results for the horizontal direction raise doubt as to the adequacy of Siipola's explanation of the correction of errors, the present study does not provide a complete test of her explanation because no data were obtained on the number of errors made.

---

[5] Siipola, *op. cit.*, 72.

[6] Scheidemann, *op. cit.*, 441.

[7] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology* (rev. ed.), 1954, 741.

[8] Siipola, *op. cit.*, 73.

It was expected that no differences would be present between the specific radii constituting each of the three general directions. This was found to be the case for the horizontal direction but not for the vertical and diagonal directions. The down vertical (180°) required less time on each response than the up vertical (0°). Among the diagonals, the only difference found was between the two up diagonals versus the two down diagonals, less time being required on each response-measure for the down diagonals. Evidently, one or more variables are operating in mirror-drawing that enable the *Ss* to perform differentially when drawing the up-down but not the right-left directions. Body orientation is a possible variable, since up-down involves movement toward or away from the body and *S* knows the actual position of her body in relation to the stimulus-figures. The possible operation of this variable is complicated by the fact that the body, which is not ordinarily visible in the mirror, is actually located in the down direction, but a down movement of the pencil appears in the mirror as an up movement. The finding previously noted that the down (vertical) directions (on the paper) were easier than the up directions means that the *Ss* did better when the pencil was perceived in the mirror as moving away from the known but not visually perceived position of the body than when the pencil was perceived as moving toward the known position of the body. This finding, unfortunately, does not provide a basis for a clear-cut statement of the operation of this variable in this complicated task.

Practice-effects. It was expected that time for the first half would be greater than for the second for each direction. The expectation was confirmed for the vertical and diagonal directions but not for the horizontal—perhaps because a reversal is present at the starting point of the vertical and diagonal directions but not of the horizontal.

The results of the analysis of data for successive blocks of two trials clearly indicate the powerful influence of practice on this task, even within the eight trials employed in this experiment. The finding that practice has greater effects for time 'out' than for time 'in' indicates that the task of correcting errors is more susceptible to improvement than is the task of drawing while inside the figure.

Summary. The experiment investigated the effect of vertical, horizontal, and diagonal directions upon the speed of mirror-drawing. Each of the 48 *Ss* drew each of eight figures once, and five measures of speed were obtained.

The major findings are: (a) For every measure, the order from least to most time is: vertical; horizontal; and diagonal. (b) For each direction, the time inside the figure is greater than that spent outside of it. (c) For each measure, a difference exists between the up and down vertical figures and between the two up diagonal and the two down diagonal figures but not between the left and right horizontal figures nor between the two left and the two right diagonal figures. (d) The time required to draw the first half of a figure (which includes the equivalent of the turn point of a regular mirror-drawing figure) is greater than that to draw the second half for the vertical and diagonal directions but not for the horizontal directions. (e) The effects of practice are clearly evident.

METHOD OF PRETRAINING AND KNOWLEDGE OF RESULTS IN PAIRED-ASSOCIATE LEARNING UNDER CONDITIONS OF REPETITION AND NON-REPETITION

By H. RAY BRACKETT and WILLIAM F. BATTIG,
University of Virginia

Following a paradigm originally introduced by Rock,[1] several recent investigators have compared the learning of a paired-associate list by the usual procedure of *repetition* (*R*), in which the same list of pairs is repeated on all trials until some criterion of performance is achieved, with a modified method of *non-repetition* (*NR*) in which all pairs not responded to correctly on any given trial are immediately removed from the list and replaced on the next trial by new pairs which *S* has never seen before. Rock and others have found no difference in performance between these two procedures,[2] which has led to the conclusion that paired-associate items are learned in all-or-none fashion in a single trial rather than building up gradually in strength through repetition. However, several more recent studies have found the *R* condition to be superior,[3] with one reported instance of significantly better performance under the *NR* condition.[4] Numerous uncontrolled variables inherent in Rock's paradigm have been suggested to account for the discrepancies in these results, and have raised serious questions about its potential usefulness as a means of testing the issue of all-or-none vs. incremental learning. This is true particularly in view of the inevitable progressive selection of easier items pos-

---

* Received for publication July 19, 1961. This research was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, and is based in part on a Master's thesis by the first author.

[1] Irvin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193.

[2] Irvin Rock and Walter Heimer, Further evidence of one-trial associative learning, this JOURNAL, 72, 1959, 1-16; Michael Wogan and R. H. Waters, The role of repetition in learning, *ibid.*, 72, 1959, 612-613; L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, *ibid*, 73, 1960, 22-40.

[3] J. C. Reed and W. D. Riach, The role of repetition and set in paired-associate learning, this JOURNAL, 73, 1960, 608-611. G. R. Lockhead, A re-evaluation of evidence for one-trial associative learning, *ibid.*, 74, 1961, 590-595; Leo Postman, Repetition and paired-associate learning, *ibid.*, 75, 1962, 372-389; B. J. Underwood, R. Rehula, and G. Keppel, Item selection in paired-associate learning, *ibid.*, 75, 1962, 353-371.

[4] Clark, Lansford, and Dallenbach, *op. cit.*, 35-38.

sible under the *NR* procedure,[5] along with the virtual impossibility of isolating the associative process to be tested for its possible all-or-none nature from other processes necessarily involved in paired-associate tasks.[6]

In the present study, attention was focused directly upon the effects of knowledge of results and of method of learning adopted by *S* on relative performance under *R* and *NR* conditions. Although in Rock's original study *S* had been informed immediately of the correctness of his response,[7] the first two reported cases of significant superiority for the *R* condition had involved no such knowledge of results.[8] These facts suggested that the continued presentation of pairs to which *E* had said 'wrong' under the *R* condition may have disturbed *S*'s subsequent performance on these pairs, an interfering effect not present under the *NR* procedure. Furthermore, *S*s under the *NR* procedure will clearly perform best by concentrating on a few pairs on each trial and ignoring completely the remainder of the list, in which case the elimination of these ignored incorrect pairs would not be expected to interfere with over-all performance. Consequently, such a 'part' method of learning should lead to superior *NR* performance and less difference from the *R* procedure than a 'whole' method by which *S*'s effort is distributed more equally over the entire list. Although two separate experiments were conducted to investigate the influence of these two factors, they were sufficiently similar so that they can be most economically described together.

METHOD

Apparatus and materials. The learning materials consisted of nonsense-syllables of 47–53% associative value,[9] from which pairs were constructed so as to avoid obvious intra-pair associations or structural similarities. Based on the results of a previous scaling procedure for ease of learning,[10] 72 non-overlapping pairs were selected from the middle range of scale-values and divided into six equivalent subsets of 12 pairs each, matched as to mean and *SD* of scale-values. These materials were photographed and presented for learning by means of an automatic slide-projector on a translucent screen located directly in front of *S*. The screen was equipped with a manually-operated shutter which enabled *E* to block off half the screen when necessary in order to present the stimulus-syllables alone for attempted recall.

---

[5] J. P. Williams, Supplementary report: A selection artifact in Rock's study of the role of repetition, *J. exp. Psychol.*, 62, 1961, 527-528; Underwood, Rheula, and Keppel, *op. cit.*, 353-371.

[6] Postman, *op. cit.*, 372-389; W. F. Battig, Paired-associate learning under simultaneous repetition and nonrepetition conditions, *J. exp. Psychol.*, 64, 1962, 87-93.

[7] For example, Rock, *op. cit.*, 186-193.

[8] Chenzoff and Gregg, *op. cit.*; Reed and Riach, *op. cit.*, 608-611.

[9] J. A. Glaze, The association value of non-sense syllables, *J. genet. Psychol.*, 35, 1928, 255-267.

[10] Battig, Scaled difficulty of nonsense-syllable pairs consisting of syllables of equal association value, *Psychol. Rep.*, 5, 1959, 126; Comparison of two methods of scaling nonsense-syllable pairs for ease of learning, *ibid.*, 6, 1960, 363-366.

Subjects. The *S*s were 160 college students who were paid for their services, divided equally among Experiments I and II. Each *S* of Experiment I was assigned unsystematically to one of 4 groups of 20 each, whereas in Experiment II 8 equivalent groups of 10 *S*s each were formed partially on the basis of performance on a pretraining task.

Design and procedure. Each *S* in all groups learned a list of 12 nonsense-syllable pairs to a criterion of one errorless trial or to a maximum of 12 trials. Each trial began with the presentation of all 12 pairs, at which time *S* pronounced both syllables of each pair, followed by a recall-series in which the first (stimulus) members of all pairs were presented alone in a different order. At this time, *S* attempted to pronounce the second (response) syllable which had previously been paired with each syllable presented. A constant 5-sec. rate of presentation was employed within both presentation- and recall-series, with a 30-sec. interval between trials and between the presentation- and recall-series within each trial.

In both experiments, half of the *S*s learned under the usual procedure of repetition (*R*), in which the same list of 12 items was presented on all trials. A condition of non-repetition (*NR*) was used for the other 40 *S*s in each experiment, in which all pairs responded to incorrectly on any given trial were removed from the list and replaced on the next trial by new pairs. Each of the 6 equivalent subsets of 12 pairs was used equally often as the *R* list and as the initial list for the *NR* group. Replacement-pairs for the *NR* group were selected from another equivalent 12-pair subset until exhausted, then from another subset, and so on until all 6 subsets had been used. Each subset also appeared with equal frequency at the various points in the sequence of replacement-pairs. In those cases where the entire pool of 72 pairs was exhausted before *S*'s final trial, subsequent replacements consisted of the eliminated incorrect pairs in the order in which they had been eliminated.

Within each experiment, half the *S*s in both the *R* and *NR* groups received immediate knowledge of results (*e.g.* were told 'right' or 'wrong' by *E* immediately following each response during the recall-series). The remaining *S*s received no such information concerning the correctness of their responses, and are therefore designated as the *NK* (no knowledge of results) group as contrasted with the *K* (knowledge of results) group.

All *S*s in Experiment I received identical, standard instructions for paired-associate learning, with an added statement that "new pairs might be shown from trial to trial" for both the *R* and *NR* groups, but received no pretraining. In contrast, all *S*s in Experiment II were given one of two types of pretraining. On the previous day, half of the *S*s within each of the four combinations of *R–NR* and *K–NK* conditions had served under a whole-pretraining condition, whereas the remaining 10 *S*s in each condition had undergone a part-pretraining procedure, both of which are described more fully in the following section. As in Experiment I, the 6 subsets of 12 nonsense-syllable pairs were used with approximately equal frequency, as the initial list for the *R* and *NR* conditions and at the various points in the sequence of replacement-pairs, for each of the various subgroups differing in knowledge of results and pretraining.

Pretraining procedure. The pretraining materials consisted of nonsense-shapes of low association-value and of high discriminability as stimuli, with two-digit numbers between 12 and 98 (excluding double numbers and numbers ending in zero) as responses. These pairs, 72 in number, had been constructed so as to be highly homogeneous in learning difficulty, and were further divided into 6 equivalent lists of 12 pairs each. Under part-pretraining, each *S* learned each of the 6 lists to a criterion of 4 correct responses. In contrast, *S*s given whole-pretraining learned only 2 of the 6 lists, each to a criterion of one errorless trial. Both types of pretraining were given during a single session of approximately 1 hr. on the day prior to the nonsense-syllable learning, and were preceded by instructions for paired-associate learning which emphasized the necessity of learning only 4 pairs of the list during part-pretraining, and of learning the entire list under whole-pretraining. Apparatus, procedure, and conditions of presentation were the same as for the nonsense-syllable pairs, except that half of the pretraining lists for each *S* were learned under the usual method of anticipation for paired-associate learning, in which each trial consisted of the initial presentation of each stimulus-shape alone, followed immediately by that shape paired with its number-response, with *S* instructed to anticipate the correct number before its appearance. For the other one or three lists of the pretraining conditions, the same method of recall was employed as was used in all cases for the nonsense-syllable pairs.[11] Through a counterbalancing procedure, half of the *S*s in each group learned their first one or three lists under the method of anticipation with the remainder using the method of recall, while the reverse was true for the remaining *S*s. Lists and combinations of lists were also counterbalanced with respect to both method and order for whole-pretrained *S*s. Both pretraining conditions corresponded to condition *NK* in that no information as to correctness of responses was given by *E*, so that the methods of anticipation and recall were comparable in this respect.

The standard instructions for paired-associate learning used in Experiment I were appropriately modified for each pretraining condition of Experiment II. Under whole-pretraining, the instructions for both pretraining and nonsense-syllable learning included at least four added statements indicating that *S* was to try to learn all pairs presented on each trial. The part-pretraining instructions, on the other hand, had specifically stated that *S* was to learn only 4 of the 12 pairs in each list. Prior to the nonsense-syllable list, these *S*s were also told that they were to learn the entire list, but that this would be less difficult if they concentrated on learning a few of the pairs on each trial, and that the purpose of the pretraining had been to give them experience with this more efficient part-method of learning. Within each pretraining condition, *S*s were assigned to the various *R–NR* and *K–NK* conditions, so that each was approximately equivalent in terms of average performance in pretraining.

RESULTS

Since fewer than half the *S*s reached the criterion of one errorless trial within 12 trials, total errors represented the basic measure of performance, for which means and *SD*s are presented in Table I. All groups under the repetition-condition (*R*), along with the non-repetition-groups given part-pretraining, are shown by Table I to be approximately equal, and markedly superior to NR groups following either whole-pretraining, or no pretraining. No effect of any kind is evident for the variable of knowledge of results (*K–NK*).

---

[11] The method and results for the whole-pretraining *S*s have been reported in detail elsewhere (W. F. Battig and H. R. Brackett, Comparison of anticipation and recall methods in paired-associate learning, *Psychol. Rep.*, 9, 1961, 59-65). Although not relevant to the present report, pretraining performance under the method of recall proved to be markedly superior to that for the method of anticipation.

Analysis of variance of the total error-scores combined for both experiments revealed significant over-all differences between *R* and *NR* conditions ($F_{1/148} = 30.71$), and pretraining conditions ($F_{2/148} = 4.58$), as well as a significant interaction of *R–NR* with pretraining ($F_{2/148} = 4.13$). Subsequent comparisons showed all *R* groups and the *NR* groups given part-pretraining to be significantly superior to those *NR* groups receiving whole-pretraining or no pretraining. These differences are clearly indicated in Table I. No other differences between groups approached significance.

TABLE I

TOTAL ERRORS UNDER EACH CONDITION

Condition	Pretraining condition					
	None (Exp. I)		Whole		Part	
	M	*SD*	*M*	*SD*	*M*	*SD*
R(*K*)	33.20	15.29	34.30	22.89	33.00	27.71
R(*NK*)	31.70	22.08	38.30	28.28	33.30	26.31
R(*K*+*NK*)	32.45	19.00	36.30	25.80	33.15	27.02
NR(*K*)	59.10	19.19	58.00	31.23	35.10	23.00
NR(*NK*)	63.55	22.72	66.40	24.71	35.80	28.11
NR(*K*+*NK*)	61.33	21.15	62.20	28.47	35.45	25.68

Although the *K–NK* effect and all interactions involving this variable fell far short of significance (all *F*s < 1), a trend-analysis demonstrated a significantly greater improvement in performance under Condition *K* over trials, as measured by total errors on each successive block of four trials ($F_{1/148} = 8.34$). Neither the slight superiority of Condition *NK* on Trials 1-4 nor that of Condition *K* on Trials 9-12 was, however, significant. The trend-analysis also showed the relative superiority of *R* over *NR* to increase significantly over trials ($F_{1/148} = 26.05$), although again only following whole-pretraining or no pretraining, as indicated by the further finding of a significant linear component of the triple interaction of *R–NR* and pretraining conditions with blocks of trials ($F_{1/148} = 4.48$).[12]

---

[12] For those *S*s reaching the criterion of one errorless trial, the total error-measures are biased downward to the extent that errors would have occurred fol-

Of the 78 *S*s reaching the criterion of one errorless trial, 50 (64%) had served under Condition *R*, yielding a X^2 of 11.02 ($p < 0.001$). These proportions, however, did not differ significantly between either pretraining or *K–NK* conditions, and separate comparisons within each pretraining condition revealed the proportion reaching criterion under Condition *R* to be significantly greater only for Experiment I (no pretraining).

DISCUSSION

From the present results, it is apparent that performance under *R* and *NR* conditions is differentially affected by whole- and part-pretraining. This finding supports the initial hypothesis that a 'part' method, in which *S* concentrates on learning only a few pairs on each trial, is optimally efficient under the *NR* procedure. By contrast, performance under Condition *R* showed no effect of either part- or whole-pretraining. As a consequence, performance under NR can be expected to vary from equivalance to marked inferiority to the *R* condition, depending on the extent to which *S*s adopt a method of part-learning. Since past studies comparing *R* and *NR* conditions have neither controlled for, nor obtained any information concerning method of learning, it can only be suggested that this factor may account for the discrepant results of these studies. Contrary to this position is the fact that specific instructions to "try to learn all pairs presented," typically given in previous studies showing no difference between *R* and *NR* conditions,[13] were omitted under the present no-pretraining condition, which nevertheless was indistinguishable in performance from the whole-pretraining condition. The present argument thus leads to the paradoxical conclusion that the *S*s of the present Experiment I, uninstructed as to learning method, had adhered more closely to a 'whole' method than those *S*s in other studies explicitly instructed to use a whole method. Still, it must be recognized that such minor instructional variations are notorious for their unreliable and inconsistent effects on learning tasks of this kind. But irrespective of the actual role of the method of learning in previous studies, the present results unequivocally demon-

---

lowing the first errorless trial. However, the magnitude of any such bias is unlikely to be substantial, and its direction and meaningfulness with respect to the problem at hand are matters of considerable uncertainty. Nevertheless, in view of this possibility, the analysis of variance was repeated for total errors on the first four trials, for which data were actually obtained from all *S*s. Except for the non-significance of the interaction of *R-NR* with pretraining ($F_{1/148} = 2.12$), the results of this analysis were virtually identical to those for total errors based on all 12 trials.

[13] Rock, *op. cit.*, 187; Clark, Lansford, and Dallenbach, *op. cit.*, 24.

strate significant differences in performance under *NR* under whole- and part-learning methods.

Since the *R* condition was virtually identical to the whole-pretraining procedure but differed considerably from part-pretraining, a relative facilitation of performance under *R* following whole-pretraining might be expected due to greater similarity to the pretraining condition. That performance under *R* was instead equivalent following whole- and part-pretraining may therefore indicate an actual over-all superiority of a part-learning method, which is not limited to the *NR* condition. However, a similar explanation of the differences in performance under *NR* following part- and whole-pretraining, in terms of greater dissimilarity of the latter to the *NR* procedure, is inconsistent with the absence of a difference between performance under *NR* following whole-pretraining and no pretraining. It is thus doubtful that specific sources of transfer from differences in formal similarity of *R* and *NR* conditions to pretraining procedures could account for the present results, particularly in view of the different learning tasks and materials employed in pretraining.

Inasmuch as replacement of incorrect pairs under the *NR* condition was found to retard learning significantly, except under conditions where *S*s have been pretrained in a part-method, the present results are clearly inconsistent with an all-or-none theory of paired-associate learning. They do not, however, preclude the possibility that a component associative phase or process may take place in all-or-none fashion within a single trial, since the present nonsense-syllable materials also required considerable response-learning and perhaps other component processes distinguishable from the associative phase of paired-associate learning.[14] Moreover, lack of control over the method of learning adopted by *S*s under *R* and *NR* conditions must now be added to the already lengthy list of objections to Rock's original paradigm in terms of its relevance to the issue of all-or-none learning. The equal distribution of *S*'s attention over all pairs of the list clearly represents an essential prerequisite to the meaningful use of a comparison between *R* and *NR* conditions with regard to the normal role of repetition in paired-associate learning.[15]

Although the initial hypothesis that performance under *R* and *NR* would be differentially affected by the knowledge of results received no support whatsoever, the further finding of no over-all difference between *K* and *NK* conditions is difficult to reconcile with the long-standing principle that complete and immediate knowledge of results leads to more rapid learning and to a higher level of performance.[16] Since knowledge of results was not completely absent under the *NK* condition (except for incorrect *NR* pairs), but merely delayed until the pairs were presented again on the next trial, these results are entirely consistent with recent findings indicating that immediate and delayed knowledge of results do not differ in effectiveness for complex human performance,[17] and can be added to the body of evidence contrary to the assumption that knowledge of results in complex tasks for humans can be treated as a simple extension of principles of reinforcement derived from studies of animal learning.[18] The demonstrated greater improvement with trials under *K* as compared with *NK* conditions, however, is consistent with an interpretation in terms of reinforcement of the effects of saying 'right' and 'wrong,' since the predominance of negative reinforcement resulting from 'wrong' responses on early trials shifts to a preponderance of 'right' positive reinforcements as trials progress. But whatever the relationship between knowledge of results and reinforcement, there is little indication that either factor played a major role under the conditions of the present study.

---

[14] For example, Postman, *op. cit.*, 372-389; Underwood, Rehula, and Keppel, *op. cit.*, 353-371; Battig, *op. cit.*, 1962, 87-93.

[15] See Postman, *op. cit.*, 372-389 for a further discussion of this point.

SUMMARY

Paired-associate learning of nonsense-syllables under a procedure of non-repetition (*NR*), in which all incorrect pairs are removed from the list and replaced by new pairs after each trial, was shown to be significantly inferior to the usual method of repetition (*R*), following no pretraining or a pretraining procedure emphasizing a 'whole' method of learning. Following part-pretraining, emphasizing the concentration on a few pairs in the list on each trial, performance under *NR* improved to a level equivalent to performance under *R*, which did not vary as a function of pretraining conditions. These findings are interpreted as contrary to a one-trial all-or-none theory of paired-associate learning, and indicate that the differential adoption of methods of whole- and part-learning may represent an important uncontrolled factor accounting for discrepancies in previous comparisons between *R* and *NR*. Knowledge of results (saying 'right' or 'wrong' following each attempted response) had only a slight effect on rate of learning, and influenced neither over-all performance nor relative proficiency under *R* and *NR* conditions.

---

[16] R. B. Ammons, Effects of knowledge and preformance: A survey and tentative theoretical formulation, *J. gen. Psychol.,* 54, 1956, 283.

[17] E. A. Bilodeau and I. McD. Bilodeau, Variation of temporal intervals among critical events in five studies of knowledge of results, *J. exp. Psychol.,* 55, 1958, 603-612.

[18] See B. F. Skinner, *Verbal Behavior,* 1957, Chap. 2.

THE INFLUENCE OF LANGUAGE AND AGE ON GESTALT LAWS OF PERCEPTION

By HANS G. FURTH and RUBEN A. MENDEZ,
Catholic University of America

A constant interest in the field of perception centers around the role of language in the reproduction of visual stimuli. Gibson maintains that changes found in the reproductions of visual forms are determined by "cues from a verbal analysis which was made of the forms during perception."[1] The role of verbal stimulation in the reconstruction of forms during their reproduction has been the topic of experimental research which originated in the studies by Wulf and by Carmichael, Hogan, and Walter.[2] On the other hand, Gestalt psychologists, without denying the influence of language, consider the organization of percepts as an autonomous activity which is perfected through maturation. As Church expresses it, "Gestalt psychology, with its theory of innate organizing factors in perception, assumes a primary experience of pure, objective objects not yet shrouded in or animated by labels, values, and meanings."[3]

The present study differs from previous attempts at determining the role of language in perception of visual forms by the use of a language-deficient population. It is common knowledge that people profoundly deaf from birth or early childhood, in general lack, to a serious degree, knowledge of and facility in using the particular language of their society. In addition, they are only minimally exposed to the impact of what may be termed the general verbal experience which may conceivably facilitate a hearing child's cognitive development. The use of deaf and hearing *Ss* within two age-groups therefore allows for the comparison among *Ss* at different levels of maturation and verbal experience.

---

* Received for publication July 24, 1961. This study was supported in part by a grant from the United States Office of Vocational Rehabilitation.

[1] J. J. Gibson, The reproduction of visually perceived forms, *J. exp. Psychol.*, 12, 1929, 1-39.

[2] Friedrich Wulf, Beiträge zur Psychologie der Gestalt: VI. Ueber die Veränderungen von Vorstellungen (Gedächtnis und Gestalt), *Psychol. Forsch.*, 1, 1921, 333-373; Leonard Carmichael, H. P. Hogan, and A. A. Walter, An experimental study of the effect of language on the reproduction of visually perceived forms, *J. exp. Psychol.*, 15, 1932, 73-86.

[3] Joseph Church, *Language and the Discovery of Reality: A Developmental Psychology of Cognition*, 1961, 4.

The view that deaf people are merely deficient in language is not, however, generally accepted. Studies of perception among the deaf led Myklebust and Brutten to conclude that auditory impairment alters visual perception.[4] As the first of these authors states, "when auditory sensations are lacking or present only to a minimal degree, the nature of his [the deaf child's] perceptions, imagery, symbols, and concepts is altered . . . development of certain types of abstract behavior is impeded."[5] On the other hand, Larr and Rosenstein found no significant differences between deaf and hearing *Ss* in perceptual tasks once the verbal limitations of the deaf were taken into consideration.[6] It is, therefore, an additional aim of the present study to provide evidence on the perceptual ability of deaf persons.

METHOD AND PROCEDURE

The figures used in this experiment were patterned after those used to illustrate Gestalt principles of perceptual organization. The purpose of using these figures was to present to the *Ss* simple materials which could be inspected and easily reproduced with minimal verbal instructions. To enable quantification of the results, the figures were presented in serial order, *i.e.* in an order accentuating the Gestalt principle in question. It was thus possible to compare, for every step, the proportions of drawings where changes had been introduced by each group of *Ss*. A sample of the figures illustrating the Gestalt principles involved is shown in Fig. 1; a description of their presentation is given below. The figures were shown on a memory-drum for 1 sec. each, allowing for immediate reproduction on a fresh sheet of paper after every exposure.

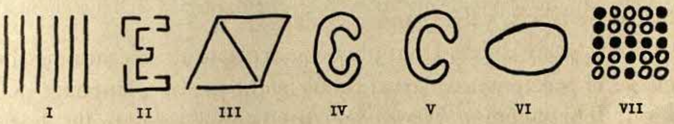


FIG. 1. SAMPLE OF STIMULI FOR THE SEVEN TASKS
(I and II represent Step 4; VII, Step 5, the remaining (III–VI) Step 1 in their respective series.)

(1) Proximity. The principle of proximity was investigated by a series of drawings in which six parallel lines were gradually separated into pairs. In the first drawing, the lines were equidistant; in those following, the space between the neighboring pairs was increased by 1/64 in. until the difference reached a ratio of 2 to 1.

---

[4] H. R. Myklebust and Milton Brutten, A study of the visual perception of deaf children, *Acta Otolaryng.* Suppl. 105, 1953, 1-126.

[5] Myklebust, *The Psychology of Deafness: Sensory Deprivation, Learning, and Adjustment,* 1960, 229.

[6] A. L. Larr, Perceptual and conceptual ability of residential school deaf children, *Exceptional Children,* 23, 1956, 63-66, 88; Joseph Rosenstein, Cognitive ability of deaf children, *J. Speech Hearing Res.,* 3, 1960, 108-119.

(2) Closure. The principle of closure was observed in the organization of a percept as a letter of the alphabet and a complete geometrical figure. *(a) Letter.* A mutilated letter 'E' was shown which required five steps in its completion. After completion of step 2, it was pointed out to all *Ss* that they were to see a letter of the alphabet. Thus, *Ss* were scored for recognition of the letter rather than closure of gaps in the reproduction. *(b)* Geometrical figure. A parallelogram with a bisecting line was presented in three steps. At Step 1, the bisecting line stopped 1/16 in. short of the angles of the figure at either end and one of the angles was open by 1/8 in.

(3) Good Form. To study the principle of good form, again a letter of the alphabet and a geometrical figure were used. *(a) Letter.* A letter *C* was presented in two series, one ascending (Steps 1 2 3) and one descending (Steps 3 2 1). On the ascending series, a protrusion in the center of the letter gradually disappeared; on the descending, the lower branch of the letter was gradually shortened. The descending series, in contrast to all other tasks, can more properly be considered a task of discrimination than of Gestalt. In other words, this series moved from a persistence in rather than an achievement of Gestalt principle. *(b) Geometrical figure.* An oval-shaped figure with a blunt end gradually rounding out was presented in three steps.

(4) Similarity. To investigate the principle of similarity, a square figure was composed of 25 small circles in which a pattern was made to emerge progressively by filling some of the circles with deeper shades of gray. To help with non-verbal instructions, a sample was shown immediately preceding the task. In the sample, the filled-in circles formed an 'X' and the *Ss* were shown that all they had to reproduce was the letter. In the experimental task, the pattern formed a figure most closely resembling the Greek letter Psi.

An effort was made to equate the *Ss* not only in terms of age (two groups with a mean age of 9 and 16 yr. and a range of 24 mo.), but also with regard to intelligence. Due to the various instruments used to measure this factor, the equation of the groups was at best quite crude. The *Ss* were chosen to constitute a normally distributed sample with a mean *IQ* of 100 for each of the four groups. Each group was also about equally divided according to sex. The deaf population in this study came from one residential school and from one day school for the deaf. In addition, the deaf *Ss* were subdivided into poor and good readers on the basis of their reading achievement-scores.

RESULTS

The results obtained appear in Table I, with symbols beneath each pair of percentages where a statistically significant difference was obtained. Statistical analysis consisted of the comparison of pairs of numbers of *Ss* by the method of binomial distribution. Since the comparison between poor and good deaf readers did not yield significant results, this distinction will not appear in the remainder of this paper.

(1) Proximity. The only significant difference occurred at Step 4, where a greater percentage of older *Ss* reproduced the lines as pairs than did the younger group.

TABLE I

Percentage of *Ss* Reproducing Drawings According to Gestalt Laws of Organization

Age (yr.)	Group	*N*	Proximity		Closure				Good form						Similarity	
					Letter E		Geom. fig.		Letter ascen.		Letter desc.		Geom. fig.			
			Step 4	Step 6	Step 4	Step 5	Step 1	Step 2	Step 1	Step 2	Step 2	Step 1	Step 1	Step 2	Step 1	Step 2
9	Deaf	32	37.5	87.5	18.8	59.4	53.1	96.9	96.9	100	96.9	78.1	100	100	21.9	46.9
9	Hearing	24	29.2	95.8	12.5	62.5	25.0*	54.2*	79.2	87.5	87.5	33.3†	79.2	91.7	58.3*	95.8*
16	Deaf	41	68.3	97.6	73.2	87.8	53.6	80.5	87.8	92.7	53.6	4.9	87.8	95.1	68.3	90.2
16	Hearing	37	45.9	97.3	27.0†	56.8*	67.6	89.2	81.1	89.2	59.4	13.6	64.9	94.6	83.8	97.3
9+16	Deaf	73	54.8	93.2	49.3	75.3	53.4	87.7	91.8	95.9	72.6	37.0	93.2	97.3	47.9	69.9
9+16	Hearing	61	39.3	96.7	21.3†	59.0	50.8	75.4	80.3	88.6	70.5	21.3*	70.5	93.4	75.4*	96.7
9	Both	56	33.9	75.0	16.1	60.7	41.1	78.6	89.3	94.6	92.8	58.9	91.1	96.4	39.3	66.1
16	Both	78	57.7*	93.6	51.3†	73.1	60.2*	84.6	84.6	91.0	56.4†	9.0†	76.9	94.9	75.6†	93.6*

Statistically significant differences (binomial distribution): * $p<0.05$, † $p<0.01$.

(2) Closure. While on other problems nearly all the *S*s eventually achieved the desired Gestalt, on Closure-Letter only 68% of the total number of *S*s succeeded in recognizing the letter E. This may have been due to the limited exposure-time. The table shows that the older *S*s recognized the letter sooner and closed the geometrical figure sooner than the younger. In addition, more 16-yr. old deaf *S*s recognized the letter than the hearing *S*s of that age, while for the geometrical figure, the younger deaf surpassed the hearing *S*s in drawing the Gestalt.

(3) Good form: (a) Letter, (i) Ascending series. The results, summarized in the table, suggest a notable equality of perceptual performance. Although the general trend points to a closer adherence to Gestalt organization by the deaf, none of the differences was statistically significant. *(ii) Descending series.* Note that Step 1 refers to the least Gestalt-like figure presented. The 9-yr. old deaf persisted in their perception of the well-formed letter longer than the hearing *S*s of the same age, although by the end of the task they perceived the distortion as frequently as the hearing *S*s. The significant difference obtained between the deaf and hearing groups in general can be attributed to this difference between the two 9-yr. old groups, since the deaf did not differ significantly from the hearing group at the older age-level. Furthermore, the combined older *S*s discriminated among the drawings sooner and more frequently than the younger group. *(b) Geometrical figure.* There were no significant differences among the groups with the geometrical figure although, as in the ascending series with the letter 'C,' the trend was toward a greater discrepancy between the younger deaf and hearing, with the older group showing an equal ability in discriminating among the drawings.

(4) Similarity. The reproduction of the pattern in this task was achieved sooner and more frequently by the older than by the younger group. As once before, the general difference favoring the hearing over the deaf group was due mainly to the performance by the younger deaf.

DISCUSSION

As evidenced by the significant differences obtained, the organization of percepts according to Gestalt principles appears to be more related to the age or maturation of the *S*s than to their specific experience with language. This generalization can best be substantiated by considering each variable separately in its relationship to the reproduction of the Gestalt figures.

With respect to language, the deaf differed from the hearing groups in three out of seven tasks. In one of these tasks (Closure-Letter), the language-deprived population recognized an incomplete letter earlier than the control group. On the figure-ground problem of Similarity, the younger deaf were less successful than the comparable hearing group. These results are apparently in agreement with Myklebust and Brutten's findings with a similar age group.[7] The present authors differ, however, from the latter's argument for a perceptual deficit in deaf children for the following two reasons: *E* reported that the deaf children encountered difficulties in understanding instructions on this problem, a difficulty which from the beginning was controlled, but only partially overcome, by the use of a sample figure. As an indication of the special difficulty of non-verbal instructions on this task, no other part required a practice problem. Secondly, the older deaf group appeared to *E* to grasp the instructions as well as the hearing group and subsequently performed at a level with their hearing peers. The third difference resulted from the one task (Good Form, Letter-Descending series) in which the persistence in, rather than the achievement of, the Gestalt principle was tested. The four problems previous to this one undoubtedly engendered a behavioral set toward completion of the Gestalt. The younger deaf group may have adhered to this set more strongly simply because deaf people, restricted chiefly to non-verbal cues, are often less secure than the hearing about the meaning of and the instructions in any problem situation. Here also, the difference obtained in the final grouping—where both age-levels were combined—can be seen to stem quite clearly from the performance of the younger deaf *S*s.

Following Werner's thought, it should be emphasized that on all perceptual tasks two tendencies may be observed: the principles of Gestalt, on the one hand, and the developmental principle of differentiation, on the other.[8] Depending on the task, these principles may work together or in opposite directions. It is noteworthy that the two problems in which the younger deaf were inferior to the hearing children call primarily for a discriminatory as opposed to a global approach.

Of all the stimuli used in the various tasks, only the lines in the Proximity series provided a relatively pure problem for the operation of Gestalt principles. The other stimuli presented possibly familiar and meaningful configurations, an additional factor influencing their recognition and reproduction. Therefore, the results on the Proximity task strongly attest to a basic equality between deaf and hearing persons in perceptual

---

[7] Myklebust and Brutten, *op. cit.*, 89.

[8] Heinz Werner, *Comparative Psychology of Mental Development*, 1948, 111ff.

Gestalt performance. If anything, the over-all trend was toward an earlier perception of the pairing-off by the deaf *S*s, both younger and older. Similarly, the results on the Closure tasks seem to indicate that lack of verbal experience does not adversely influence the closure of these Gestalten, even though one was a highly verbal symbol.

These observations lead to the conclusion language-deprived *S*s, while not differing significantly from the verbally trained groups, more consistently organize their percepts according to the principles of Gestalt organization. Following the developmental formulations of such authors as Werner and Church, one may reason that this tendency on the part of the deaf *S*s is a manifestation of their reliance on preverbal modes of perceptual organization. As was pointed out by Furth, the deaf person's superior performance on certain problems may be due to his less sophisticated approach to the situation, in which case the hearing person's greater store of available verbal mediators may be actually a hindrance.[9] While emphasis on Gestalt organization or misunderstanding of instructions led the younger deaf to a faulty perception and reproduction of some figures which required breaking up the totality of a pattern, the older *S*s, regardless of their facility with language, performed on a comparable level.

The fact that Gestalt organization, however, is not to be regarded simply as a pre-verbal or pre-adolescent type of perceptual approach is best demonstrated by comparing the results obtained from the two age-groups. On the tasks illustrating two of the four Gestalt principles tested (Proximity and Closure), the older *S*s succeeded better precisely because they applied the Gestalt principle sooner and more frequently.

In conclusion, more significant differences were obtained by comparing *S*s of differing ages of maturational level than by the comparison of *S*s at different levels of language-experience. Thus, the importante of age, *i.e.* general, non-verbal experience, to Gestalt laws of perception may be said to be relatively greater than the influence of the acquisition of specific language. Regarding perceptual ability of the deaf *S*s, their reproduction of Gestalt forms in general gave little evidence of any basic difference from that of the hearing *S*s.

SUMMARY

To test for the influence of age and language on some Gestalt laws in visual perception, seven different figures were presented, each in a serial

---

[9] H. G. Furth, The influence of language on the development of concept formation in deaf children, *J. abn. soc. Psychol.*, 63, 1961, 386-389.

arrangement with successive drawings gradually accentuating the Gestalt principle in question. Two groups of hearing *S*s with a mean age of 9 and 16 yr., respectively, were compared with two similar groups of deaf *S*s, hence language-deficient youngsters. Results in general revealed that age was a discriminating variable, while verbal experience was not. These findings were interpreted as suggesting that Gestalt principles were formed and developed relatively independently from the acquisition of language.

SERIAL LEARNING AS A FUNCTION OF PRIOR PAIRED-ASSOCIATE TRAINING

By ROBERT K. YOUNG, EDMUND W. MILAUCKAS,
and JAMES D. BRYAN, University of Texas

Variables related to transfer of training have received a great deal of attention in experiments concerned with verbal learning. Although some studies have investigated transfer of serial habits to subsequent serial lists, *e.g.* Irion, the major portion of studies of transfer-of-training has dealt with the relationships between paired-associate *(PA)* training and subsequent *PA*-learning.[1] Underwood, for example, identifies four paradigms which relate interlist similarity of stimulus and response to direction and amount of transfer from one *PA*-list to another.[2]

One immediately obvious reason for the extensive use of *PA*-learning in studies of transfer is that stimulus- and response-variables, such as interlist and intralist similarity, may be manipulated independently in *PA*-paradigms, while such is not the case in serial (*Sr*) learning. Since each item must serve both as a stimulus and as a response in *Sr*-learning, development of paradigms of transfer analogous to those used in *PA*-studies appears to be difficult.

An approach to the problem of developing paradigms of transfer in *Sr*-learning might be made through the study of relationships between *PA*-learning and subsequent *Sr*-learning. Two studies have investigated such transfer, and in both the paradigm of transfer was such that the items and *S–R* relationships in the *PA*- and *Sr*-lists were identical.[3] Positive transfer, relative to a control, obtained from the *PA*- to the *Sr*-learning.

The present study extends previous investigations of transfer from *PA*- to *Sr*-learning. Three paradigms of transfer were constructed. The relationships between *PA*- and *Sr*-lists were such that either positive, negative, or

---

* Received for publication August 4, 1961. The present study was supported in part by Grant G-4850 from the National Science Foundation and in part by the Graduate School of The University of Texas.

[1] A. L. Irion, Retroactive inhibition as a function of the relative serial position of the original and interpolated items. *J. exp. Psychol.*, 36, 1946, 262-270.

[2] B. J. Underwood, *Experimental Psychology: An introduction*, 1949, 303.

[3] Ernest Primoff, Backward and forward associations as an organized act in serial and in paired associate learning, *J. Psychol.*, 5, 1938, 375-395; R. K. Young, A comparison of two methods of learning serial associations, this JOURNAL, 72, 1959, 554-559.

zero transfer was expected. The paradigm of positive transfer was the same as that used in the studies by Primoff and Young. The paradigm of negative transfer employed the same items in both *PA*- and *Sr*-lists but with the *S–R* relationships being randomly changed from the *PA*- to the *Sr*-list. This technique of re-pairing is similar to the paradigm of negative transfer employed by Porter and Duncan in *PA*-learning.[4] The condition of zero transfer was simply a control in which the items used in the *PA*-list were different from those used in the *Sr*-list.

Underwood, in a *PA*-study, found an interaction between degree of original learning and paradigm of transfer.[5] That is, transfer increased in his condition of positive transfer as degree of original learning increased, while transfer in his condition of negative transfer was not related to degree of learning of the first list. These same relationships may hold in transfer from *PA*- to *Sr*-learning and, for this reason, degree of *PA*-training has been varied in the present study.

The *PA*-lists employed by Primoff and by Young differed from typical *PA*-lists in that the items in the list served both as a stimulus for one pair and as a response for another.[6] The same construction of the *PA*-lists was used in the present study, hence at the start of *Sr*-learning each item in the series would have served previously as a stimulus and as a response in the *PA*-learning. Thus the apparent function of each item was the same in both lists.

Method. Design. The *S*s (108) serving in this study were taken from classes in introductory psychology and randomly assigned to groups. Each *S* was given training on one of three *PA*-lists and then learned a common *Sr*-list. Three groups of 18 *S*s each received 15 trials on a *PA*-list, and three other groups, also composed of 18 *S*s, each received 30 trials on a *PA*-list. Following *PA*-training, all *S*s learned the same *Sr*-list to a criterion of one perfect recitation. The three *PA*-lists employed in the study were so arranged that either positive (Group P), negative (Group N), or zero (Group 0) transfer would be expected to occur from the *PA* to the serial list. The experimental design may then be described as a 2 × 3 factorial design with the two rows representing degree of *PA*-training and with the three columns representing the paradigm of transfer employed.

Lists. Two *PA*-lists of eight pairs each were developed from an *Sr*-list of nine items. These nine items were adjectives of low similarity taken from Melton.[7] The *PA*-list used in Group P was so constructed that the *S–R* arrangements in the *PA*-

---

[4] L. W. Porter and C. P. Duncan, Negative transfer in verbal learning, *J. exp. Psychol.*, 46, 1953, 61-64.

[5] Underwood, Associative transfer in verbal learning as a function of response similarity and degree of first-list learning, *ibid.*, 42, 1951, 44-53.

[6] Primoff, *op cit.*, 379; Young, *op cit.*, 555.

[7] E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 517-567.

and *Sr*-lists were identical. Thus, in Group P any correct association formed during *PA*-learning would be compatible (*i.e.* correct) with the ordering of the subsequent *Sr*-items. The *PA*-list used in Group N was so constructed that the *S–R* arrangements in the *PA*-list would be incompatible (*i.e.* incorrect) with the ordering of the subsequent serial items. A third *PA*-list, also containing eight pairs and composed of nine dissimilar adjectives, was used in the condition of control. These items were as unrelated to the items in the *Sr*-list as careful inspection would permit.

With the exception of the first and last, each item in all three *PA*-lists served both as a stimulus and as a response. The pairs were assembled in lists of paired associates, each with three orders of presentation and with the restriction that pairs with items in common would not follow one another. The *Sr*-list and the three *PA*-lists used in the study are presented in Table I.

TABLE I

LISTS USED IN THE PRESENT STUDY

PA-Lists						*Sr*-List
Group P		Group O		Group N		All Groups
SOLVENT	UNWELL	PUBLIC	WEARY	RUSTIC	JOYOUS	ALIKE
ENTIRE	RUSTIC	ORAL	SWOLLEN	TABOO	ALIKE	ICY
ICY	JOYOUS	AWKWARD	LUSTROUS	ICY	UNWELL	JOYOUS
UNWELL	CLEVER	WEARY	EAGER	JOYOUS	CLEVER	SOLVENT
ALIKE	ICY	DIRECT	AWKWARD	ENTIRE	ICY	UNWELL
CLEVER	ENTIRE	EAGER	ORAL	CLEVER	TABOO	CLEVER
JOYOUS	SOLVENT	LUSTROUS	PUBLIC	UNWELL	RUSTIC	ENTIRE
RUSTIC	TABOO	SWOLLEN	MODERN	ALIKE	SOLVENT	RUSTIC
						TABOO

TABLE II

MEAN CORRECT RESPONSES FOR EACH GROUP ON LAST THREE TRIALS OF *PA*-LEARNING

	Groups					
PA-Trials	P		O		N	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
15	12.54	4.53	10.78	3.99	12.22	5.08
30	17.22	4.81	18.22	4.01	16.44	5.32

Procedure. The lists were presented by a projector with a 2:2-sec. rate for the *PA*-lists and a 2-sec. rate for the *Sr*-list. The intertrial interval was 6 sec. for both *PA*- and *Sr*-lists. The time between learning of the *PA*- and *Sr*-lists was approximately 3 min. to allow *E* to read instructions for *Sr*-learning and to change filmstrips. Each *S* served in only one session with no practice being given.

Results: Equality of PA-*lists*. Performance of the six groups on the last three trials of *PA*-training was compared. The means and standard deviations of numbers of correct responses for the six groups are presented in Table II. In the analysis of the data presented in Table II only degree of training was significant ($F = 37.02$, with 1 and 102 *df*, $P < 0.001$).

At the end of training the *S*s in the groups receiving 15 trials were giving about four correct responses per trial and the *S*s in the groups receiving 30 trials were giving about six correct responses per trial. The data presented in Table II tend to indicate that the three *PA*-lists did not differ in difficulty and that *PA*-learning was related to degree of *PA*-training.

Serial learning. Means of trials to learn the *Sr*-list as a function of condition of transfer and degree of *PA*-training are presented in Table III. The variances of the six conditions were tested by the $F_{\text{max.}}$ test and were found to be heterogeneous ($F_{\text{max.}} = 5.04$, with $K = 6$ and $N = 18$, $P < 0.05$).[8] Observation of Table III indicates that the *SD* of Group N in the 30-trial condition was considerably larger than the other *SD*s and is responsible for the significant heterogeneity. The data have not been transformed, and the analysis performed on these data, indicates that rate of *Sr*-learning was related to paradigm of transfer ($F = 21.19$, with 2 and 102 *df*, $P < 0.001$). Orthogonal comparisons between the three conditions of transfer were made.[9] With two degrees of freedom associated with condition of transfer two such comparisons can be made. The P *vs* N, O comparison indicates that learning in Group P was faster than in Groups O and N when considered together ($F = 22.73$, with 1 and 102 *df*, $P < 0.001$), while the N *vs* O comparison indicates that learning in Group N was slower than in Group O ($F = 19.64$, with 1 and 102 *df*, $P < 0.001$). No other *F*s were significant in the analysis. So that differences between conditions could be described more fully, two additional analyses were made between Groups P and O. With reference to the groups receiving 15 trials of *PA*-training, the *F* obtained between Groups P and O was found to equal 1.30, which does not approach significance, while in the analysis of the groups receiving 30 trials the *F* between Groups P and O was equal to 4.87 (with 1 and 34 *df*, $P < 0.05$). In the over-all analysis, the inter-

TABLE III

MEAN TRIALS TO CRITERION FOR THE SERIAL-LIST IN EACH OF THE SIX GROUPS

	Groups					
	P		O		N	
PA-Trials	Mean	*SD*	Mean	*SD*	Mean	*SD*
15	6.56	3.67	7.83	3.03	11.44	2.95
30	5.28	2.95	7.33	2.63	11.44	5.66

---

[8] Helen Walker and Joseph Lev, *Statistical Inference,* 1953.

[9] V. L. Senders, *Measurement and Statistics,* 1958.

action between degree of *PA*-training and condition of transfer did not approach significance nor did Group P vary as a function of degree of *PA*-training ($F = 1.33$). Degree of *PA*-training does not appear to be a significant variable in the present experiment ($F < 1.00$).

Means of overt errors per trial were also analyzed. The means for each group are presented in Table IV. An analysis of variance revealed only one significant *F*. Condition of transfer was related to means of overt errors per trial ($F = 4.42$, with 2 and 102 *df*, $P < 0.05$). Comparisons between the three conditions of transfer indicate that the major source of variance was in Group P. The *F* between Groups P and O was found to equal 3.90 with 3.94 ($df = 1$ and 102) required for significance at the 0.05 level.

TABLE IV

MEANS OF OVERT ERRORS PER TRIAL IN EACH OF THE SIX GROUPS DURING *Sr*-LEARNING

PA-Trials	Groups					
	P		O		N	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
15	.77	.45	.84	.42	1.18	.40
30	.62	.47	1.01	.57	.89	.57

DISCUSSION

The results of the present experiment indicate that rate of *Sr*-learning is a function of prior *PA*-training. When associations are formed in *PA*-training which are incompatible with the ordering of items in a subsequent *Sr*-list, negative transfer will occur. In addition, the results of the present study tend to confirm the results of previous research in which positive transfer was found from *PA*- to *Sr*-learning when the items and *S–R* arrangements in both *PA*- and *Sr*-lists were identical.[10]

Although previous research has suggested that the nominal stimulus in *Sr*-learning may not always be the effective stimulus, it is apparent that associations formed during *PA*-training do transfer to an *Sr*-task.[11] Thus, while the effective stimulus may be some other aspect of the situation in *Sr*-learning (perhaps ordinal position), rate of *Sr*-learning is a function of associations formed between the nominal stimulus and some response in a prior *PA*-task. In Group P, when an association is formed between *S* and

---

[10] Primoff, *op. cit.*, 375-396; Young, *op. cit.*, 554-559.

[11] Young, The stimulus in serial learning, this JOURNAL, 74, 1961, 517-528; Tests of three hypotheses about the effective stimulus in serial learning, *J. exp. Psychol.*, 63, 1962, 307-313.

R in *PA*-learning, it will be 'used' to *S*'s advantage in *Sr*-learning whether it is "natural" to do so in usual *Sr*-learning. So, too, will an association between *S* and *R* be 'used' in *Sr*-learning in Group N despite the fact that such use will retard learning. In any event, it is apparent that associations formed in prior *PA*-learning influence rate of *Sr*-learning, even though such associations may not be formed usually in *Sr*-learning.

The magnitude of positive transfer in Group P after 30 *PA*-trials was not as great as might be expected if *S* were simply re-learning the same associations in the *Sr*-list as were learned in the *PA*-task. Aside from the typical postcriterion-drop which is observed even within the same list, several factors may have contributed to the less-than-perfect transfer from the *PA*- to the *Sr*-list.[12] Initially, it may be noted that degree of *PA*-learning was not extremely high. The average *S* was giving about six correct responses per trial at the end of *PA*-training. In addition, the change in rate of responding may have reduced the apparent transfer from the *PA*- to the *Sr*-list. *S* had to change his response-set from a response every 4 sec. to a response every 2 sec. While both of these factors may have contributed to the lack of complete transfer, they may not be critical since transfer is not perfect even when *S*s have been trained to a criterion of one perfect recitation and have had practice in *Sr*-learning.[13] Finally, contextual changes may account for the absence of complete transfer. For example, Greenspoon and Ranyard found in a study of retroactive inhibition that retention of original learning is less if relearning takes place in a context different from that of the original learning.[14] In the present experiment the change from *PA*- to *Sr*-learning may have been sufficient to reduce the amount of apparent transfer.

The results of the present experiment tend to support the results of Underwood's experiment despite the fact that the interaction between degree of *PA*-training and transfer paradigm was not significant.[15] As in Underwood's experiment (a) negative transfer in Group N did not increase as a function of degree of training and (b) positive transfer in Group P did increase as a function of degree of training. This increase was not in itself significant and the only justification for this assertion is that the positive transfer obtained following 15 trials of *PA*-learning was not significant while the transfer obtained after 30 trials was.

---

[12] Underwood, A graphical description of rote learning. *Psychol. Rev.*, 64, 1957, 119-122.

[13] Young, *op. cit.*, 554-559.

[14] Joel Greenspoon and Redge Ranyard, Stimulus conditions and retroactive inhibition, *J. exp. Psychol.*, 53, 1957, 55-59.

[15] Underwood, *op. cit.*, 44-53.

As in other rote learning experiments, the error-data are not directly related to rate of learning.[16] The analysis of errors shows that means of overt errors per trial decrease in Group P but do not differ in Groups O and N. One possible explanation of these data would be that the incorrect associations learned by Group N in the *PA*-training were relatively easy to discriminate in the subsequent *Sr*-learning and were not elicited.

Summary. Six groups of *S*s learned a paired-associate (*PA*) list and then a serial (*Sr*) list. The variables investigated in the study were degree of *PA*-training (either 15 or 30 trials) and paradigm of transfer (either positive, negative, or zero). The major results of the experiment were:

(1) Rate of serial learning was directly related to paradigm of transfer employed.

(2) Degree of *PA*-training was not an effective variable.

(3) Fewer overt errors occurred in Group P than in Groups O or N during serial learning.

---

[16] Underwood, Studies of distributed practice: IX. Learning and retention of paired adjectives as a function of intralist similarity. *J. exp. Psychol.*, 45, 1953, 143-419; Young, Paired associate learning when the same items occur as stimuli and responses, *J. exp. Psychol.*, 61, 1961, 315-318.

MULTIDIMENSIONAL ANALYSES OF CONFUSIONS OF ENGLISH CONSONANTS

By KELLOGG V. WILSON, University of Nebraska

In this paper, multidimensional scaling methods are used to identify the cues used in the perception of speech. The methods are applied to data previously collected by Miller and Nicely.[1] These authors treated their data by partitioning the results in accordance with the accepted articulatory characteristics of speech-sounds. They confirm the view that the effective acoustic cues are related to the manner in which the sounds are produced. An alternative approach, used by Liberman and his associates, has been to test the discriminations that *S*s could make among synthetically produced speech sounds.[2] The use of scaling methods, and factor analyses, provides an independent check on the correctness of the analytic schemes used in the earlier studies.

Data used in analysis. The experiment of Miller and Nicely consisted of testing for the perceptual confusions among various English consonants spoken before the vowel phoneme *a* as in f*a*ther. Five women were presented orally with a randomly selected list of consonants and were instructed to identify the consonants they heard. The task of reading the consonants was rotated among the *S*s and the consonants were presented at various signal-to-noise ratios and with varying degrees of acoustic filtering. Two sets of consonants were used; one contained the 6 English stopped consonants and another set of 16 contained these 6 consonants plus 10 others. The data were reported in the form of matrices of frequencies indicating how often each stimulus was reported as each of the possible consonants.

The matrices for a signal-to-noise ratio of —12 db. for a 200 to 6500 cps. channel were used in the analyses.[3] These matrices were selected since the signal-to-noise ratio was low enough to permit a considerable amount of differentiation between the consonants, yet high enough to permit application of the analytic techniques described below.

---

* Received for publication June 1, 1960. The first analyses in this study, carried out at Duke University with the assistance of Mr. John Ost, were supported by National Science Foundation Grant G-4842. The final analyses were made at Michigan State University. The author is indebted to Dr. Charles Wrigley, Mr. Phillip Snett, and to the Laboratory Staff of the Digital Computer of that institution for assistance.

[1] G. A. Miller and Patricia E. Nicely, An analysis of perceptual confusions among some English consonants, *J. acoust. Soc. Amer.*, 27, 1955, 338-352.

[2] A summary of this work and references to previous studies are contained in A. M. Liberman, Some results of research on speech perception, *J. acoust. Soc. Amer.*, 27, 1957, 117-123.

[3] Miller and Nicely, *op. cit.*, Tables II and XVIII.

Method of analysis. Multidimensional analysis when used to scale stimuli consists of two steps. In the first step, a measure of distance is computed between each pair of stimuli. These measures increase with increasing dissimilarity between the stimuli and imply an arrangement of the stimuli in a multidimensional space (nearly always Euclidean), much as the tables of distances between cities found on road maps imply the arrangement of the cities on the map proper. The second stage of the analysis consists of calculating a more economical set of coördinates for the stimuli which are consistent with the interpoint distances. Factor analytic methods are used in these solutions and the dimensions of these orderings correspond to the dimensions of the stimuli which are sufficient to describe their perceived similarity. These coördinates describe a 'map' of the stimuli in multidimensional space.[4]

The aspect of multidimensional stimulus-scaling which is most open to controversy is the choice of a distance-measure, since these measures are generally based on arbitrary and often unprovable assumptions. Accordingly, two distance-measures based on very different kinds of assumptions have been selected for use in this analysis. The formulas for computing these measures from the matrices of judgment-frequencies are given below.

Shepard's distance measure: $$d_{ij} = .5 \log \left[\frac{f(i/i) \times f(j/j)}{f(i/j) \times f(j/i)} \right]$$

Wilson's distance measure: $$d_{ij} = \sqrt{1 - \frac{f(i/j) + f(j/i)}{f(i/i) + f(j/j)}}$$

where d_{ij} is the distance between stimuli i and j; $f(j/i)$ is the frequency with which stimlulus i is identified as j; $f(i/j)$ is the frequency with which stimulus j is identified as i; and $f(i/i)$ and $f(j/j)$ are the frequencies of correct identifications for stimuli i and j, respectively.

Shepard's measure is derived from a complex and extensive model of stimulus- and response-generalization.[5] Wilson's measure is not based on such an elaborate rationale, but is based on the assumption that independent differences in cue should have additive effects in reducing the proportion of confusional errors.[6] Both measures do, however, have two features in common. First, both formulae yield symmetric distance-matrices, *i.e.* $d_{ij} = d_{ji}$. Secondly, the distance between a stimulus and itself is always zero, *i.e.* $d_{ii} = 0$. The first feature may well obscure some important asymmetries in the matrix of $f(i/j)$ values and the second seems somewhat unreasonable for this application, since the random white noise presented with the consonant sounds has a masking effect which could be regarded as creating a random distribution of effects such that the average distance between repeated consonants would be non-zero. Both of these possibly inappropriate properties seem, however, to be required of a distance-measure and we have no more appropriate analytic model to use instead.[7]

---

[4] A review of these methods is contained in W. G. Torgerson, *Theory and Methods of Scaling*, 1958, Chap. 11.

[5] The rationale for Shepard's measure may be found in R. N. Shepard, Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space, *Psychometrika*, 22, 1957, 325-345, and Stimulus and response generalization: Deduction of the generalization gradient from a trace model, *Psychol. Rev.*, 65, 1958, 242-256. Empirical tests of the measure may be found in Shepard, Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space, *J. exp. Psychol.*, 55, 1958, 509-523. The formula for Shepard's measure given above is not the same as that given by him, but it is an equivalent form which is more convenient for computational purposes.

[6] The distance-measure proposed is simply the square root of one minus the ratio of frequnecy of confusion for a pair to frequency of correct identification. If adding an independent cue-difference to a pair reduces this proportion by a constant amount, then the squared inter-stimulus distance could be represented by a sum of effects of independent cues. The Pythagorean theorem can be shown to hold from this result, since independent cue-differences are represented on orthogonal axes.

Once the distance-measures have been computed, the next step in multidimensional stimulus-scaling is to compute a B-matrix from the interstimulus distance-measures. In computing the B-matrix, it is necessary to select an arbitrary origin for the coördinate system. This origin was placed at the centroid of the configuration of points.[8] The B-matrix is then factor analyzed to yield an A-matrix whcih contains the coördinates of the stimuli. The Principal-Axes-Method of factor analysis was used without subsequent rotation, since this method has two very substantial advantages. First, this solution places the orthogonal reference axes on or parallel to the lines of symmetry of the configuration of points when the origin is placed at the centroid—a solution which is quite easily interpreted since it is much like a 'simple structure' solution and which avoids the arbitrariness of rotation. Secondly, the latent roots of the B-matrix are obtained which indicates the relative importance of dimensions and their validity in ordering the stimuli in a Euclidean space—the larger positive latent roots corresponding to the more important dimensions and negative latent roots indicating invalid dimensions.[9]

ANALYSIS OF CONFUSIONS OF SIX ENGLISH STOPPED CONSONANTS

Results. The results of the analysis of the confusions of the six stopped consonants are given in Table I, which contains the more important dimensions from the A-matrices. Thus in the column headed "I" are given the projections of each of the stimuli on the axis having the largest positive root, and in Column II, the projections on an orthogonal axis having the next largest root. The first three dimensions or factors for both measures are given. For Shepard's measure, only the dimensions with sizable positive latent roots are given—one of the remaining three possible dimensions had a very small positive latent root, and two turned out to have fairly large negative roots. For Wilson's measure, again the dimensions corresponding to large positive latent roots are given—two of the remaining three dimensions corresponding to small positive latent roots and one corresponding to a very small negative latent root.

---

[7] See Shepard, *op. cit.,* 1957, 333, for a discussion of the metric axioms for a distance-measure.

[8] Torgerson (*op. cit.,* 254-259) gives the formula in Equation [16] which was used to compute the B-matrix from the inter-stimulus distances.

[9] Actually, an error-free ordering of stimuli in *k*-dimensional Euclidean space will yield a B-matrix with *k* positive latent roots and the remaining latent roots will be zero. In practice, however, empirically obtained B-matrics ordinarily do contain some negative latent roots. See Torgerson, *op. cit.,* 255-256.

Discussion. The results of these two analyses are obviously somewhat different, but they both show some agreement with the articulatory and acoustic cues which distinguish these six consonants.[10] For both measures, Factor I corresponds to voicing, the positive loadings being associated with unvoiced sounds and the negative loadings being associated with voiced sounds. It is interesting that /b/ has a somewhat intermediate status for both measures. Factor II for Shepard's measure seems to correspond to placement of the voiced consonants, with a positive loading corresponding to front placement. There is a single inversion in the order, of /k/ and /t/, but even this might be corrected by a rotation of II away from III by about 30°. The values for the unvoiced consonants are somewhat less than the voiced consonants. This is not unreasonable, since the white noise should more effectively mask unvoiced than voiced sounds. Factor II for Wilson's measure serves principally to distinguish /b/ from the remainder. Except for the reversal in sign, this is not inconsistent with Factor II for Shepard's measure.

TABLE I

SOLUTIONS FOR COÖRDINATES OF ENGLISH STOPPED CONSONANTS

	Shepard's measure			Wilson's measure		
	I	II	III	I	II	III
/p/	+.71	+.39	+.05	+.40	+.02	−.32
/t/	+.60	−.28	−.23	+.39	+.07	+.30
/k/	+.67	−.19	+.24	+.44	+.08	+.01
/b/	−.49	+.41	−.26	−.23	−.60	+.02
/d/	−.79	+.08	+.32	−.50	+.22	−.05
/g/	−.71	−.41	−.11	−.49	+.22	+.03
Latent root	+2.71	+0.62	+0.30	+1.06	+0.48	+0.20

For Factor III, Wilson's measure corresponds to placement for the unvoiced sounds and, again, the position of /t/ and /k/ is reversed even more markedly than for Shepard's measure. This may be due to masking by the white noise, as noted above, or to Miller and Nicely's hypothesis that /t/ is distinguished from /p/ and /k/ by its higher frequency of noise. For Shepard's measure, Factor III is somewhat difficult to interpret, but it seems to involve some sort of interaction between voicing and placement—the middle placement having a negative loading when it is unvoiced and a positive loading when voiced.

---

[10] See Miller and Nicely, *op. cit.*, 346-347.

ANALYSIS OF CONFUSIONS OF SIXTEEN ENGLISH CONSONANTS

Results. The results of the analyses of the confusions of 16 English consonants are given in Table II below which contains the more important dimensions of the A-matrices. Again, the values in the column headed "I" represent the projection of the respective sound on the first Principal Axis, and under "II" if on the second axis, and so on.

The first four dimensions or factors for both measures are given in Table II, along with the corresponding latent roots. For Shepard's measure, 4 of the remaining 12 possible factors are associated with small positive latent roots, while the remaining 8 are associated with negative latent roots, 2 of which are less than minus one. For Wilson's measure, 10 of the remaining 12 factors corresponded to small positive latent roots and 2 to small negative latent roots.

Discussion. As was the case for the stopped consonants, the two measures result in somewhat different solutions but we again find some agreement with the articulatory cues which distinguish these consonants. For both measures, Factor I is again clearly a voicing factor with the positive loadings corresponding to the unvoiced sounds. The rank orders of the loadings do not, however, agree as well as for the previous analyses. For Factor II, we find the greatest amount of disagreement between the two measures. For Shepard's measure, Factor II gives sizable positive loadings to the longer duration consonants, /s/, /ʃ/, /z/, /ʒ/, and to /d/ and /g/ and sizeable negative loadings to the two nasal consonants /m/, /n/. Obviously, this is a complex and not easily interpreted dimension. For Wilson's measure, Factor II gives sizable positive loadings to the two nasal consonants and near-zero loadings to the others and so is obviously a nasality factor. For both measures, Factor III seems to correspond to placement with negative loadings corresponding to front placement for Shepard's measure, except for /m/, and positive loadings corresponding to front placement for Wilson's measure. Both measures, however, share exceptions to this pattern. First, the unvoiced stopped consonants, /p/, /t/, /k/ are not adequately placed, but this can be explained by the relatively low energy-level of these sounds relative to the background noise. Secondly, both measures indicate fairly sizable differences between /θ/ and /s/ and between /ð/ and /z/, whereas there are no substantial differences in placement for these sounds. These pairs do differ, however, in duration so perhaps there is an interaction between placement and duration for Factor III. For Shepard's measure, Factor IV corresponds to whatever distinguishes between /b/ and /ʒ/ at one extreme and /s/ and /v/ at the other. For Wilson's measure, Factor IV gives sizable positive loadings to the four unvoiced fricatives /f/, /θ/, /s/, /ʃ/, sizable negative loadings to the three unvoiced stop consonants /p/, /t/, /k/ and near-zero loadings to the others. This is obviously an affrication factor for the unvoiced consonants.

METHODOLOGICAL CONSIDERATIONS

Since both the methods of multidimensional scaling analysis and the cues distinguishing English phonemes from each other are relatively unfamiliar, a few methodological comments would seem to be in order.

It should be appreciated that the factor solutions used here are unrotated orthogonal ones and so may not show the same agreement with the articulatory or acoustic cues as solutions which are rotated to maximize the degree of agreement, particularly if an oblique rotation is used. There is some justification for an oblique rotation for the matrices of 16 consonants, since two of the articulatory features are correlated in that set—all of the sounds of long duration used are also fricatives. Rotation, however, was not attempted.

TABLE II

SOLUTIONS FOR COORDINATES OF SIXTEEN ENGLISH CONSONANTS

	Shepard's measure				Wilson's measure			
	I	II	III	IV	I	II	III	IV
/p/	+.52	−.11	+.35	+.03	+.50	−.14	−.07	−.29
/t/	+.93	−.33	+.23	+.06	+.51	−.13	−.08	−.25
/k/	+.88	−.40	+.14	+.38	+.49	−.14	−.05	−.28
/f/	+.79	−.20	−.61	−.09	+.26	−.04	+.26	+.25
/θ/	+.54	−.08	−.22	−.20	+.30	−.07	+.17	+.28
/s/	+.66	+.45	+.01	+.15	+.21	−.02	−.02	+.42
/ʃ/	+.72	+.43	+.49	−.46	+.17	+.06	−.12	+.34
/b/	−.46	−.19	−.53	+.39	−.29	−.02	+.37	−.13
/d/	−.44	+.35	−.17	−.22	−.37	−.17	−.14	−.00
/g/	−.59	+.30	+.19	−.07	−.45	−.19	−.18	−.05
/v/	−.58	+.09	−.47	−.40	−.27	+.01	+.34	−.08
/ð/	−.40	−.12	−.54	+.17	−.40	−.14	+.26	−.13
/z/	−.34	+.69	−.16	+.13	−.35	−.18	−.21	+.04
/ʒ/	−.60	+.56	+.51	+.37	−.21	−.02	−.45	+.05
/m/	−.68	−.90	+.13	−.21	−.05	+.58	+.01	−.11
/n/	−.96	−.64	+.64	−.02	−.04	+.61	−.11	−.08
Latent Root	+6.89	+2.88	+2.42	+1.03	+1.81	+0.90	+0.76	+0.73

Another question is which of the two measures used is the better? If we assume that the distance-measures should order stimuli in multidimensional Euclidean space, then Wilson's measure has proven more satisfactory, since the B-matrices associated with this measure had fewer and smaller latent roots. There is no reason, however, to believe that this will always be the case. Shepard's measure has the more adequate theoretical rationale, but it is quite possible that there are situations in which this rationale is invalid. More complete studies may lead to a set of analytic methods, each appropriate to a different underlying process.[11]

SUMMARY

Multidimensional methods of stimulus-scaling were applied to matrices of the confusions in the identification of English consonants under noisy conditions. Two measures of inter-stimulus distance were used in the analyses that were carried out for the 6 English stopped consonants and for a larger set of 16 consonants. While the results for the two distance-measures differed somewhat from each other, they both showed some agreement with the articulatory and acoustic differences between these sounds.

---

[11] Torgerson (*op. cit.*, 292-293) suggests that different distance-models may be more appropriate when the stimulus dimensions are "obvious and compelling" to the *S*s than when they are not; C. H. Coombs, On the use of inconsistency of preferences in psychological measurement, *J. exp. Psychol.*, 55, 1958, 1-7, contains an empirical study which indicates that the relationship of preferential choice-inconsistency to psychological distance depends on a condition Coombs calls "laterality."

CROSS-ADAPTATION TO THE ALIPHATIC ALCOHOLS

By TRYGG ENGEN, Brown University

It has been suggested that the study of intensive and qualitative properties of cross-adaptation (or co-adaptation) is a useful approach to the problem of olfactory stimulation and stimulus-classification.[1] The problem is to measure the similarity of odors on the basis of the extent to which previous exposure to one odor influences the detectability of another. Such results would be of value in correlating physiochemical properties and odor.

An homologous series of alcohols provides desirable stimulus-material for the present experiments for it minimizes chemical differences and specifies physical properties of stimuli. A group of investigators have recently described both a method for preparing sniff-bottles and results indicating that psychophysical suprathreshold results on intensity are correlated with the chain-length of the alcohol.[2] There is then the possibility of relating the effects of cross-adaptation to changes in known physical characteristics of the stimuli, although the results on homologous series are by no means clear-cut.[3]

The main purposes of this investigation were to determine (1) whether previous exposure to one alcohol would change *O*'s ability to detect another alcohol, and (2) if so, whether the magnitude of this effect depends upon the similarity of the two alcohols in chain-length.

---

* Received for publication July 28, 1961. This research was supported by Grant G7603 from the National Science Foundation. The author is indebted to Doris Stern Donovan, Stephen W. Klapper, and Donald F. Mathews, who served as research assistants during the various phases of this project.

[1] E. B. Tichener, *A Text-book of Psychology*, 1910, 1261; R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 316-317; F. A. Geldard, *The Human Senses*, 1953, 285-288.

[2] L. H. Beck, Lawrence Kruger, and Paul Calabresi, Observations on olfactory intensity: I. Training procedures, methods, and data for two aliphatic homologous series, *Ann. N. Y. Acad. Sci.*, 58, 1954, 225-238; Lawrence Kruger, A. N. Feldzamen, and W. R. Miles, A scale for measuring supra-threshold intensity, this JOURNAL, 68, 1955, 117-123; Comparative olfactory intensities of the aliphatic alcohols in man, this JOURNAL, 68, 1955, 386-395.

[3] Kruger, Feldzamen, and Miles, *op. cit.*, this JOURNAL, 68, 1955, 391-392; D. Ottoson, Studies on the relationship between olfactory stimulating effectiveness and physio-chemical properties, *Acta. Physiol. Scand.*, 43, 1958, 167-181.

GENERAL METHOD

The method used in these experiments has been described earlier,[4] and only essential details will be described here. The data reported were obtained from college juniors and seniors who had previous experience in sniffing test-tubes. They were run individually in groups of 4 or 5 per experiment and they were paid for their services. The experimental laboratory, without windows, is air-conditioned, its temperature maintained at 70-72°F. and humidity at 50%. The seven alcohols used in the experiments were C_3, 1-propanol; C_4, 1-butanol; C_5, 1-pentanol; C_6, 1-hexanol; C_7, 1 heptanol; C_{10}, 1-decanol; and C_{11}, 1-undecanol. With Mallinkrodt's Benzyl Benzoate (USP XIII, 3872) as the dilutent, a geometric ($\log_2$) concentration dilution-series was prepared for each alcohol with a pipette technique, described in detail by Beck, Kruger, and Calabresi,[5] in 10 × 75-mm. pyrex test-tubes. A number of tubes containing 1 cc. of pure benzyl benzoate also were prepared. Thus, there was 1 cc. of clear liquid in each test-tube, and great caution was used in checking test-tubes for other identifying marks.

EXPERIMENT I

The first experiment measured thresholds of four *O*s for C_4, C_5, and C_6 (test-stimuli) following previous exposure either to C_3 or C_7 (adapting stimuli).

Procedure. To maintain the same conditions throughout the experiment, benzyl benzoate, the relatively odorless dilutent, was used as a pseudo-adapting stimulus in measuring the absolute threshold for the three test-stimuli as baseline for the adaptation-effects. There were then nine different conditions for each of the four *O*s formed by the combination of each of the three test-stimuli with each of the three adapting odors. The nine conditions were presented in a systematic sequence designed to eliminate series-effects. Each *O* required about four weeks to make all the observations.

Thresholds were measured with the ascending method of limits. *E* presented the adapting and test-stimuli together to *O* on a small rack. Immediately following the sniff of the adapting stimulus, *O* sniffed the test-stimulus and indicated whether or not he could detect the odor he had been exposed to previously at suprathreshold concentration. *O* rested for 2 min. between each pair of adapting and test-stimuli to prevent a possible cumulative effect of the adapting stimulus. At least two consecutive affirmative answers were required before the series was terminated. *O* sniffed the full concentration (100%) of the test-stimulus at the beginning of each ascending series, again followed by a 2-min. rest before the presentation of the first pair of adapting and test-stimuli. A 10-min. rest followed each ascending series to avoid both general and olfactory fatigue. *O* completed 12 such series for each of the nine conditions. In each series two 'blanks,' *i.e.* test-tubes containing only the dilutent, were introduced randomly in the series of concentrated dilutions as check experiments.

---

[4] Trygg Engen, Effect of practice and instruction on olfactory thresholds, *Percept. mot. Skills*, 10, 1960, 195-198.

[5] Beck, Kruger, and Calabresi, *op. cit.*, 226-228.

Results. For each ascending series, threshold is defined as the midpoint between the concentration-dilution values representing the last negative and first positive observation of the presence of the alcohol-odor. Each *O's* threshold for a condition is taken as the mean $\log_2$ concentration-dilution value for the 12 ascending series. The average thresholds of four *Os* for the nine conditions are presented in a bar graph in Fig. 1. The effect of cross-adaptation among aliphatic alcohols is plain, for without exception the thresholds for the three test-alcohols, C_4, C_5, and C_6, are higher when measured after the presentation of another alcohol, C_3 or C_7, than after the benzyl benzoate dilutent (B). There are no individual exceptions.

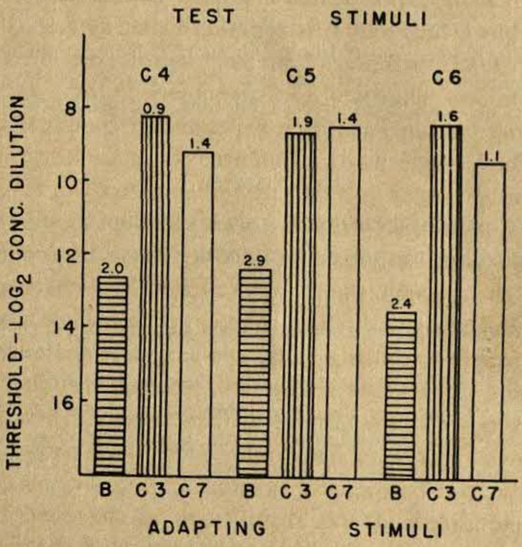


FIG. 1. THE EFFECT OF ADAPTING STUMULI C_3, C_7, AND BENZYL BENZOATE ON THE THRESHOLD LOG_2 CONCENTRATED DILUTION FOR C_4, C_5, AND C_6

The number at top of each bar is Fig. 1 is the *SD* for each condition in $\log_2$ units of the concentrated dilutions. It can be seen that the *SDs* are consistently larger for thresholds following the presentation of B than following either C_3 or C_7. This heterogeneity of variance suggests that the $\log_2$ units of the concentrated dilutions becomes subjectively smaller with weaker concentrations, but this scaling problem is of no particular interest here. The *SDs* obtained for the 12 thresholds of each *O* for each condition show the same pattern. That the *Os* were cautious is indicated by the fact that all four of them together judged only 6 of the 864 check experiments (blanks) as alcohol- (or test-) odors.

With respect to the second purpose of this investigation, there is no support in Fig. 1 for the hypothesis that the magnitude of the cross-adaptation-effect depends upon the similarity of adapting and test-stimuli in chain-length. One might expect C_3 to have the greatest effect on the threshold for C_4, next C_5, and least for C_6, while the reverse effect should be obtained for C_7.

This hypothesis was investigated in further experiments, but with variations in variables which might have nullified the effect of chain-length in Experiment I. First, the order of presentation of experimental conditions was changed to a sequence which was random with the restriction that one ascending series be run from each experimental condition per day. The purpose of this change was to eliminate the possibility that day-to-day changes might be greater than the effect relating to chain length. Secondly, the adapting stimuli of Experiment I were all 100% concentrations of the alcohols and not equal in subjective intensity. Since this could override any effect due to chain-length, the subjective intensity of the adapting stimuli were matched. Thirdly, a different selection of alcohols was used to maximize differences in chain-length; C_3 and C_{11} were the adapting stimuli, and C_4, C_7, and C_{10} were the test-stimuli, yielding chain-length differences of 1 and 7 compared with 1 and 3 in Experiment I. These manipulations did not change the results significantly from those shown in Fig. 1. It also should be noted that these experiments revealed no correlation between absolute threshold and chain-length as suggested by the suprathreshold intensities studied by Kruger, Feldzamen, and Miles.[6]

EXPERIMENT II

An interest in a possible difference between threshold and suprathreshold intensities prompted Experiment II. Experiments in cross-adaptation generally have been done on threshold intensities, but judgment of higher intensities might eliminate some of the problems related to threshold-measurement and possibly provide a better test of the hypothesis regarding chain-length.[7] In any case, the relationship between threshold and suprathreshold subjective intensities is not known.

Procedure. Three paid *O*s served in this experiment. The alcohols used were C_4, C_7, C_{10}, and C_{11}. In the first part of the experiment, the *O*s made intensity-matches of these four alcohols.[8] This produced psychophysically-equal intensities as near 100% concentration-dilution as possible. In the second part of the experiment, the *O*s were trained to identify each of the four alcohols without error by label when presented singly as in the method of absolute judgment.[9] Next followed the experiment proper. Each of the four alcohols served both as an adapting and as a test-stimulus; thus, 16 pairs of stimuli were presented, including pairs of identical stimuli. Each pair was presented 10 times in a different random sequence for each *O*. In this case, cross-adaptation was measured as reduction from 100% correct identification. In other respects, the procedure was similar to that of Experiment I with respect to rest-periods, refamiliarization with the four odors before each session, and so forth.

---

[6] Kruger, Feldzamen, and Miles, *op. cit.*, this JOURNAL, 68, 1955, 387-389.
[7] Engen, *op. cit.*, 195-198.

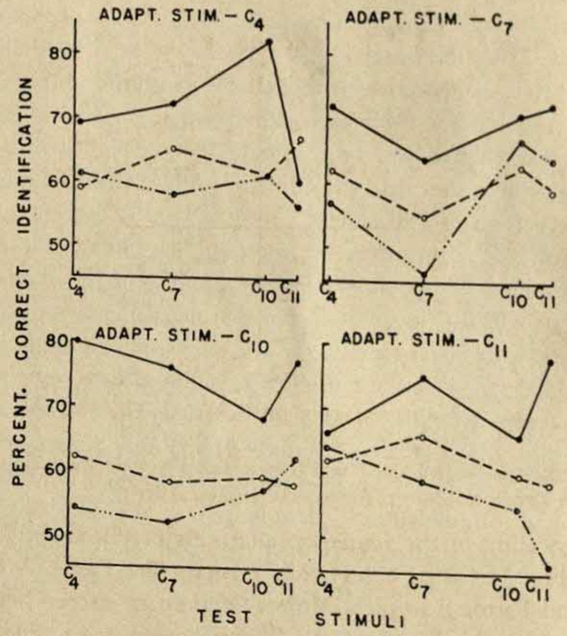


FIG. 2. CROSS-ADAPTATION AMONG C_4 C_7, C_{10}, AND C_{11} IN DETECTION OF SUPRATHRESHOLD CONCENTRATED DILUTIONS

Results. The results are presented in Fig. 2. Cross-adaptation was produced as indicated by reduction of percentage of correct identification from 100%. Without exception there was cross-adaptation, but again the magnitude of the effect did not vary systematically with the degree of similarity in chain-length.

---

[8] Kruger, Feldzamen, and Miles, *op. cit.*, this JOURNAL, 68, 1955, 386-387.

[9] Trygg Engen and Carl Pfaffmann, Absolute judgments of odor quality, *J. exp. Psychol.*, 59, 1960, 214-219.

EXPERIMENT III

The final experiment investigated the dimensionality or psychological similarity of these alcohols with the complete method of triads developed by Torgerson.[10]

Method. Forty graduate students and staff members of the Department of Psychology judged all possible triads plus rotations of C_3, C_4, C_7, C_{10}, and C_{11}. For each of the 30 triads, *O* judged which two alcohols were most similar, generally following the same procedure as for pairs of stimuli above.

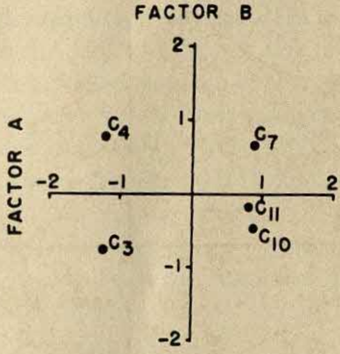


FIG. 3. FACTOR MATRIX OF C_3, C_4, C_7, C_{10}, AND C_{11}

Results. Scaling of the similarity-judgments produced the factor-matrix shown in Fig. 3. Casual observation suggests that Factor A is related to intensity and Factor B to pleasantness. Of primary interest here, however, is the fact that chain-length clearly does not yield one simple psychological dimension.

SUMMARY AND CONCLUSION

Definite cross-adaptation was obtained for the aliphatic alcohols both at threshold and at suprathreshold intensities. This finding lends support to the use of the cross-adaptation method, in agreement with recent published results.[11] Cross-adaptation as reported by *O* does not, however, appear to vary systematically with the degree of physical similarity of the stimuli as defined by carbon chain-length of alcohols. Olfactory research on homologous series represents an attempt to find a simple physical explanation for odor. The present results lend support to the belief of Pfaffmann that no single stimulus-dimension is likely to account for the complex olfactory system.[12] Multidimensional analysis of the homologous alcohols indicates a complex psychophysical relationship with two psychological dimensions. Perhaps it would be more fruitful to consider cross-adaptation on the basis of scaled psychological similarity. That work remains to be done.

---

[10] W. S. Torgerson, *Theory and Methods of Scaling*, 1958, 263-280.

[11] G. H. Cheesman and Stella Mayne, The influence of adaptation on absolute threshold measurements for olfactory stimuli. *Quart J. exp. Psychol.*, 5, 1953, 22-30; G. H. Cheesman and M. J. Townsend, Further experiments on olfactory thresholds of pure chemical substances, using the 'sniff-bottle method,' *ibid.*, 8, 1956, 8-14.

---

[12] Carl Pfaffmann, Taste and smell, *Ann. Rev. Psychol.*, P. R. Farnsworth (ed.), 7, 1956, 391-408.

CONTEXTUAL EFFECTS IN JUDGMENT AS A FUNCTION OF RESTRICTIONS IN RESPONSE-LANGUAGE

By SAMUEL FILLENBAUM, University of North Carolina

Changes in context have been found consistently to affect judgments made by the method of single stimuli. In the typical case, *S* is asked to use a situationally relative and arbitrarily limited set of categories, as in judging each of a set of weights in terms of a five-category scale ranging from 1 (very light) to 5 (very heavy); the introduction of a new stimulus which falls outside the original distribution, *e.g.* a much lighter weight, then leads to shifts in the judgments of the series-stimuli away from the category assigned to the new stimulus.[1] As has been suggested previously,[2] such shifts in judgment may reflect actual changes in the appearance of the stimuli, or they may be nothing but 'semantic' effects which depend on limitations in the response-language used by *S*.[3]

To permit a decision between these alternative interpretations, the present experiment was planned. The *S*s rated a series of rectangles on a slimness-broadness dimension, some having been given the usual fixed-scale instructions, others instructed to estimate ratio of height to base, with the distribution of stimuli manipulated in ways known to lead to predictable context-effects under the former instructions. From the semantic viewpoint, context-effects would not be expected under the latter instructions, which permitted the *S*s to deal with every stimulus-object in absolute terms.

METHOD

Experimental design. The various conditions of the experiment were studied under category-judgment instructions, requiring *S* to make his ratings of the slimness-broadness of the rectangles on a five-category scale, and also under ratio-judgment instructions, requiring *S* to make his ratings of the rectangles by estimating the ratio of height to base for each to the nearest 1/10. In Condition 1, the control condition, all seven stimuli appeared equally often. In the other conditions, the

---

* Received for publication August 7, 1961. The author is indebted to Dr. R. D. Bock for carrying out the analysis of data on the local computer.

[1] W. A. Hunt and John Volkmann, The anchoring of an affective scale, this JOURNAL, 49, 1937, 88-92; H. R. McGarvey, Anchoring effects in the absolute judgment of verbal materials, *Arch. Psychol.*, 39, 1943, No. 281, 1-86.

[2] D. T. Campbell, N. A. Lewis, and W. A. Hunt, Context effects with judgmental language that is absolute, extensive and extra-experimentally anchored, *J. exp. Psychol.*, 55, 1958, 220-228.

[3] This is the problem discussed by Stevens when he distinguishes between adaptation as a result of changes in excitability and judgmental relativity as a function of changes in the modulus of judgment. See S. S. Stevens, Adaptation-level vs. the relativity of judgment, this JOURNAL, 71, 1958, 633-646.

stimulus-distribution was manipulated. In Condition 2, the distribution was skewed to the slim end, the three slimmest rectangles appearing twice as often as the others. In Condition 3, the distribution was skewed to the broad end, the three broadest rectangles appearing twice as often as the others. When the range was extended by adding a stimulus at the slim end (Condition 4), this slim anchor-stimulus appeared twice as often as the series-stimuli. When the range was extended at the broad end (Condition 5), the broad anchor-stimulus appeared twice as often as the series-stimuli. When the range was extended at both ends (Condition 6), all nine stimuli appeared equally often. The anchor-stimuli were presented for judgment just as were the other stimuli.

On the basis of previous work with context-effects, one should expect that skewing of the distribution to the slim end (Condition 2) and extension of the range at the slim end (Condition 4) should cause the rectangles to be rated as somewhat broader, while skewing of the distribution to the broad end (Condition 3) and extension of the range at the broad end (Condition 5) should cause the rectangles to be rated as somewhat slimmer than in the control condition. Extension of the stimulus-range at both ends (Condition 6) should push the rating of the series-stimuli closer together, with the result that the slimmer rectangles should be rated broader, and the broader ones slimmer, than in the control condition, with attendant reduction in the slope of the response-function. The purpose of this study was to see whether such context-effects were specific to categorical judgment, or whether they still would occur with ratio-judgment.

Materials. Each rectangle was mounted on a 31 x 56 cm. white card which was displayed before a large dark screen for about 2 sec. The height of the rectangles was kept constant at 25 cm., and shape was varied by changing the breadth. The seven rectangles making up the stimulus-series were spaced at equal logarithmic steps. The ratio of height to base for the slimmest rectangle was 5.00:1; for the broadest rectangle this ratio was 2.67:1. The breadths of the stimuli were as follows: 5.00, 5.56, 6.17, 6.87, 7.62, 8.45, and 9.36 cm. The ratio of height to base for the slim anchor-stimulus was 6.15:1, and its breadth was 4.05 cm.; the ratio of height to base for the broad anchor-stimulus was 2.25:1, and its breadth was 11.11 cm.

Instructions and procedure. The *S*s were given sheets on which to record their responses, and the following instructions then were read twice to *S*s who were to work under categorical judgments.

> This is an experiment in shape-perception—in the perception of the shape of rectangles. You will be shown each rectangle briefly, and I wish you to judge it on the following five-point scale: 1, very slim; 2, slim; 3, medium; 4, broad; 5, very broad. Look at the rectangle, make your judgment quickly and record it on the sheet. Use all the categories of judgment.

For *S*s who were to work under ratio-judgments, the following instructions were read twice:

> This is an experiment in shape perception—in the perception of the shape of rectangles. You will be shown each rectangle briefly, and I want you to judge its slimness-broadness by estimating the ratio of height to breadth—how many times it is as tall as it is broad, how many times does the base go into the height. Estimate the ratio of height to breadth to the nearest 1/10th or 0.1; *e.g.* if you saw a figure you considered 2⅓ times as high as broad you would put down 2.3; if

> you saw a figure you thought 3 7/10 times as tall as broad you would put down 3.7; and so forth. Always estimate the ratio of height to base to the nearest 0.1. Look at each rectangle, make your estimate quickly and record it on the sheet.

Stimuli were presented in random order within blocks of 7 (Condition 1), blocks of 10 (Conditions 2 and 3) and blocks of 9 (Conditions 4, 5, and 6). The *S*s made about 140 ratings, and approximately the last 100 of these constituted the data of this study. The earlier judgments were not scored in order to permit *S*s to stabilize their responses.

Subjects. The *S*s were 308 University of California undergraduates who participated in the experiment to satisfy a class requirement. They were tested in small groups, usually three or four *S*s, and assigned at random to the various conditions. The *N* for each condition is given in Tables I and II.

RESULTS

The mean category-number assigned each rectangle under the categorical instructions is given in Table I. The mean judgment of the ratio of height to base for each rectangle under ratio instructions is given in Table II.

TABLE I

MEAN CATEGORY-NUMBER ASSIGNED EACH RECTANGLE

Condition	Anchor slim	Rectangle number 1	2	3	4	5	6	7	Anchor broad	*N*
1		1.30	1.76	2.13	2.59	3.09	3.58	4.22		40
2		1.28	2.00	2.46	2.82	3.50	3.91	4.49		22
3		1.16	1.67	1.93	2.42	2.84	3.41	4.12		19
4	1.23	2.02	2.32	2.62	3.11	3.48	3.98	4.52		21
5		1.18	1.62	2.00	2.31	2.70	3.00	3.37	4.37	21
6	1.06	1.89	2.01	2.21	2.45	2.84	3.08	3.39	4.33	21

TABLE II

MEAN JUDGMENT OF RATIO OF HEIGHT TO BASE FOR EACH RECTANGLE

Condition	Anchor slim	Rectangle number 1	2	3	4	5	6	7	Anchor broad	*N*
1		5.16	4.79	4.52	4.27	3.98	3.78	3.52		49
2		5.03	4.65	4.45	4.22	4.00	3.79	3.58		25
3		5.25	4.75	4.42	4.14	3.86	3.58	3.29		21
4	5.60	4.84	4.47	4.23	4.00	3.76	3.55	3.30		24
5		4.87	4.56	4.25	4.06	3.80	3.59	3.38	2.72	24
6	5.64	4.91	4.64	4.36	4.16	3.88	3.67	3.50	2.90	21

The raw data were broken down into their mean, linear, and quadratic components. First, a general over-all analysis was made to assess the differences among the various conditions under categorical instructions and the differences among the various conditions under ratio-instructions. The multivariate criterion used (Roy's largest-root criterion) simultaneously takes into account mean, linear, and quadratic effects, and the possible correlation among these.[4] Under categorical instructions, the statistic was significant beyond the 1% level, while under ratio instructions, the statistic did not quite reach significance at the 5% level. Further analysis for the mean, linear, and quadratic effects separately obviously was in order for the data obtained under categorical instructions. While one might perhaps have dispensed with any further tests on the data obtained under ratio-instructions, such tests were made nevertheless, since the largest-root criterion employed above is a conservative one.

A trend analysis was performed on the data using the method of orthogonal polynomials,[5] and the results of this analysis are summarized in Table III. Under categorical instructions, there were significant differences between conditions with regard to level or mean, with regard to slope, and with regard to curvature. Under ratio-instructions, only the linear or slope-term reaches significance at the 5% level. From the analyses reported thus far then, it seems clear that manipulation of the shape of the stimulus-distribution has considerable effect on response under categorical instructions and little, if any, effect under ratio-instructions.

TABLE III

TREND ANALYSIS OF RESPONSES UNDER CATEGORICAL AND RATIO INSTRUCTIONS

Source	Categorical instructions			Ratio instructions		
	df	*MS*	*F*	*df*	*MS*	*F*
Mean						
Between conditions	5	1.4754	28.16†	5	0.0412	<1
within conditions	138	0.0524		158	0.0604	
Linear						
Between conditions	5	2.4928	44.60†	5	0.2388	2.57*
Within conditions	138	0.0559		158	0.0928	
Quadratic						
Between conditions	5	0.6963	10.76†	5	0.0435	<1
Within conditions	138	0.0647		158	0.0866	

* $p<0.05$. † $p<0.01$.

Next, the difference between the control condition and each of the experimental conditions was tested separately, confidence limits based upon Scheffé's method for multiple comparisons being used.[6] First, consider the results with regard to mean or level of response. Under categorical instructions, Condition 1 (control) differed significantly beyond the 5% level from Condition 2 (distribution skewed to slim end), from Condition 4 (range extended at slim end), and from Condition 5 (range extended at broad end); all these differences were in the predicted direction. There was no significant difference between Condition 1 and Condition 6 (range extended at both ends), and none had been anticipated with regard to over-all level of respose. It had been predicted that, in Condition 3 (distribution skewed to broad end), the stimuli would be judged slimmer than in Condition 1 (control). While the differences in judgment were in the expected direction (see Table I), they did not reach significance. The Scheffé procedure, based as it is upon *all* possible comparisons, is a conservative or stringent one for this situation where only the comparisons between the control group and the various experimental groups are of interest. The difference between Conditions 1 and 3 therefore was reëvaluated using Dunnett's test, which is based upon control and experimental comparisons only.[7] For significance at the 5% level, on a one tailed test, a t of 2.26 was required in this situation but the t obtained was only 2.16. While this particular difference does not quite reach significance, the general tenor of the results for the mean or level-effect under categorical instructions is clearly in line with expectation in terms of known context-effects. As noted above, there were, with ratio-instructions, no significant differences between conditions with regard to the mean or level term.

---

[4] L. V. Jones, Some illustrations of psychological experiments designed for multivariate statistical analysis, Univ. of North Carolina, *Psychometric Laboratory Reports*, No. 28, 1960; D. L. Heck, Charts of some upper percentage points of the distribution of the largest characteristic root, *Amer. math. Statist.*, 31, 1960, 625-642.

[5] A. L. Edwards, *Experimental Design in Psychological Research*, 1960, 224-253; D. A. Grant, Analysis-of-variance tests in the analysis and comparison of curves, *Psychol. Bull.* 53, 1956, 141-154.

For the linear effect, the only prediction made was that the slope of the response-function in Condition 6 (range extended at both ends) would be less than that in Condition 1 (control). With categorical instructions, there was indeed a significant difference ($p < 0.05$) between the two conditions, and the difference was in the predicted direction. While there was no significant difference in slope between Condition 1 and Conditions 2, 3, and 4, there was a significant difference between Conditions 1 and 5 (range extended at broad end), the slope of the function in the latter condition being smaller. It is difficult to account for this finding. Comparing Conditions 5 and 6, it was found that the slope of the function in the latter condition was significantly smaller ($p < 0.05$); thus the slope of the response-function in Condition 6 is significantly smaller than that in any other condition. With the partial exception of Condition 5, the results for the slope-effect under categorical instructions were consistent with expectation. The linear or slope-effect also was evaluated under ratio-instructions by testing the difference in slope between the control condition and each of the experimental conditions. None of these differences reached significance at the 5% level.

No prediction had been made as to differences between conditions with regard to curvature or the quadratic effect. With categorical instructions a significant difference ($p < 0.05$) was found between Conditions 1 and 2, and between Conditions 1 and 5. No explanation is offered here as to why skewing the distribution to the slim end (Condition 2) or extending the range at the broad end (Condition 5) should have led to differences in curvature, or as to why, if there were such effects there were no comparable effects of skewing the distribution to the broad end (Condition 3) or extending the range at the slim end (Condition 4). Under ratio instructions, as noted earlier, there were no significant differences between conditions with regard to the quadratic effect.

---

[6] H. Scheffé, A method for judging all contrasts in the analysis of variance, *Biometrika,* 40, 1953, 87-104.

[7] C. W. Dunnett, A multiple comparison procedure for comparing several treatments with a control, *J. Amer. statist. Ass.,* 50, 1955, 1096-1121.

In general, considering the mean (level) and linear (slope) effects, the results under categorical instructions are about as predicted, and exhibit the customary contextual effects. With ratio-instructions, such effects disappear. These findings suggest that some of the contextual effects obtained when the method of single stimuli is employed may be semantic effects specific to the situationally relative and restricted response-language imposed on *S*.

The evidence on this issue in the literature is sparse. An experiment dealing with the pitch of tones judged in terms of a piano-keyboard showed a contextual effect after a rather substantial shift in the stimulus-distribution even when *S*s did use a judgmental language which was "absolute, extensive, and extra-experimentally anchored."[8] In an experiment in which the *S*s were required to judge the slimness-broadness of a set of rectangles, extension of the stimulus-range led to differences in judgment which were of borderline significance even though *S*s were permitted free choice as to the number of categories in the judgment scale.[9] In another experiment, skewing of the stimulus-distribution did not lead to significant shifts in the judgment of the series-stimuli, although it did have a significant effect upon the location of the neutral level.[10] Finally, in a recent study by Krantz and Campbell, some evidence was presented that suggests the reality both of semantic and perceptual effects.[11]

Perhaps some of these results can be reconciled if it is recalled that, in this experiment, absolute judgment *was* possible under ratio-instructions, since the presentation of each stimulus provided full information. This was not the case in other experiments on shape-judgment cited above since, even with free choice as to the number of categories in the judgmental scale, the assignment of a particular rectangle to a given category could be made only with reference to the assignment of the other rectangles to the other categories. A similar comment holds with regard to the study by Campbell, Lewis, and Hunt, insofar as none of their *S*s had perfect pitch.[12]

---

[8] Campbell, Lewis and Hunt, *op. cit.*, 220-228.

[9] Samuel Fillenbaum, The effect of a remote anchor upon judgment with salient within-series stimulus-object and with a free choice of scale, this JOURNAL, 74, 1961, 602-606.

[10] Fillenbaum, The effect of distributional skewing upon judgment with free choice of scale, this JOURNAL, 73, 1960, 132-136.

[11] D. L. Krantz and D. T. Campbell, Separating perceptual and linguistic effects of context shifts upon absolute judgments, *J. exp. Psychol.*, 62, 1961, 35-42.

SUMMARY

Some of the contextual effects found when *S*s are required to judge each of a set of items by the method of single stimuli were examined in an attempt to determine whether these effects should be regarded as semantic effects specific to the restricted response-language imposed on *S*, or whether the shifts in judgment resulting from changes in context may reflect actual changes in the appearance of the stimuli.

The *S*s rated rectangles on a slimness-broadness dimension. Those given categorical instructions were required to use a fixed, arbitrary, fives-category scale, while those given ratio-instructions made their ratings by estimating the ratio of height to base for each rectangle. With categorical instructions, the usual context-effects were obtained, while, with ratio-instructions, manipulation of the stimulus-distribution had no systematic effect on the ratings. In circumstances where absolute judgment was truly possible, with complete information for every judgment available in each single-stimulus presentation, the usual context-effects were not obtained. The findings thus suggest that at least some context-effects in judgment may be semantic effects specific to the situationally relative and arbitrarily limited language with which *S* is required to express his judgments.

---

[12] Campbell, Lewis, and Hunt, *op. cit.*, 220-228.

THE EFFECT OF RATE OF PRESENTATION, SUBSTITUTION, AND MODE OF RESPONSE IN PAIRED-ASSOCIATE LEARNING

By LEE W. GREGG, ANDREW P. CHENZOFF, and KENNETH R. LAUGHERY,
Carnegie Institute of Technology

The meaning of the phrase "to form an association" serves as a focus for the experiments to be reported here. Inferences about the processes that enable *S* to learn paired-associates range from analogies to the classical conditioning paradigm to assumptions of highly cognitive mediating acts.[1] In the conditioning analogy, there are, presumably, incremental increases in stimulus-response strength over successive repetitions of the pairs. Our own biases tend toward the view that most performance changes in the human adult depend on cognitive processes utilizing information previously acquired. The data presented by Rock were therefore of particular interest to us.[2]

It was Rock's contention that an association between the elements of a pair is formed in a single trial. Repetition is not necessary for the formation of associations, but does assume importance in that only a few associations may be formed on any one trial. Hence, to learn a long list of items, the list must be repeated.

Two experiments were carried out using the general procedure employed by Rock. In Experiment I, the *Ss* were required to spell the nonsense-syllables, instead of pronouncing them, at two different rates of presentation. In Experiment II, a direct comparison of the spell vs. pronounce modes of response was made at a single rate of presentation identical to that used by Rock. The purpose of these procedural changes was to determine the generality of the phenomenon reported by Rock. Specifically, it was hypothesized that in a more difficult task Rock's conclusions will not be supported.

---

* Received for publication August 7, 1961. This study was supported in part by a grant by the Ford Foundation to the Graduate School of Industrial Administration.

[1] A. E. Goss, C. H. Morgan, and S. J. Golin, Paired-associate learning as a function of percentage of occurrence of response members (reinforcement), *J. exp. Psychol.*, 57, 1959, 96-104; Irvin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193.

[2] Rock, *op. cit.*, 186-193.

EXPERIMENT I

Subjects. Eighty students, 57 men and 23 women, undergraduate students in psychology, participated in the experiment as volunteers. Four groups of 20 *Ss* each were obtained by assigning *Ss* to the various combinations in random replications.

Apparatus. The apparatus consisted of a large, black panel with a flashed glass screen, 6 in. high and 9.5 in. wide, mounted in the center at *S*'s approximate eye-level. An automatic slide projector with a 500-w. lamp was used to project nonsense-syllables from behind the screen. Control of the rate of presentation was accomplished by means of an electronic interval-timer.

Materials. Eighty 2 × 2-in. slides were prepared, each with a different pair of CVC nonsense-syllables selected from Glaze's lists of 47% and 53% associative value. The pairs of syllables were so constructed that no letter appeared more than once in a given pair. In addition, pairs that seemed to provide obvious associations were rejected.

The 80 pairs were assigned to a fixed order and divided into 10 lists of 8 pairs each. Each of the lists was assigned to two *Ss* in each of the four groups. The *Ss* learning under the control conditions received the same list throughout learning. As the *Ss* in the substitution conditions required new pairs, these were taken in order from the ones following the initially assigned list.

Procedure. Rate of presentation at two levels was the variable orthogonal to the control and substitution conditions. The rapid rate of presentation was 2-sec. exposure followed by a 3-sec. interval; that is, 5 sec. between the exposure of successive slides. For the slow rate the slides were on for 3.5 sec. and off for 4.5 sec. After a delay of 15 sec., the same slides were projected in the same order and at the same rate, but with the second syllable of the pair blocked from *S*'s view. The *Ss* were instructed to spell aloud both the first and the second syllables on these test-trials.

The interval between a recall-test and the next learning trial was approximately 30 sec. During this time, the order of the three slides was changed. The first slide was placed in Position 3, the second in Position 7, the third in Position 8, etc., until the slides were in the order: 3, 7, 8, 6, 2, 1, 4, 5. For the substitution-groups, the syllable-pairs missed on the previous test were replaced by new ones before the slides were rearranged.

Trials, consisting of a learning phase and recall-test, were continued until: (a) *S* recalled perfectly a list of eight syllable-pairs; or (b) *E* ran out of new pairs of syllables to substitute; or (c) a total of 16 trials had been given.

Results. Fig. 1 shows the trial-by-trial learning curves for each of the groups for the first 10 trials. At the tenth trial, 8 of the 40 *Ss* in the substitution-groups had learned so few pairs that the pool of 80 syllable-pairs was exhausted and the sessions were terminated. Of the 20 *Ss* in the substitution-group with the 2-sec. presentation-rate, none reached the criterion of one perfect recitation within 10 trials. For the 3.5-sec. presentation-rate and the substitution-conditions, 2 *Ss* out of 20 learned to this criterion within 10 trials. In the control groups, the numbers of *Ss* reaching the criterion in 10 trials were 6 and 12 for the 2-sec. and 3.5-sec. rates, respectively.

An analysis of the total number of correct responses in the first 10 trials confirms the apparent degradation in performance produced by the substitution-conditions. The means for the four groups were: 16.1 for the 2-sec. substitution-group; 17.8 for the 3.5-sec. substitution-group; 31.2 for the 2-sec. control group; and 44.2 for the 3.5-sec. control group. The main effects of Substitution-Conditions and Presentation-Rate are significant.

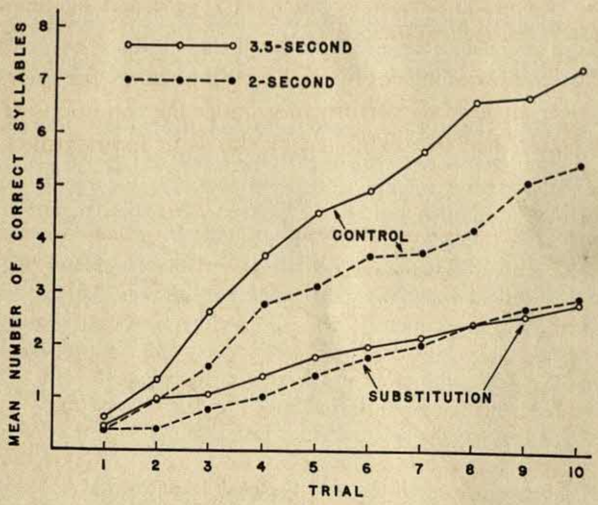


FIG. 1. ACQUISITION-CURVES FOR EXPERIMENT I
Mean number of correct responses during the first 10 acquisition trials for the control and substitution-groups with presentation rate as parameter. The *Ss* spelled the syllables.

The interaction of these variables fails to reach significance. For $p < 0.05$, the required F with 1 and 19 *df* is 4.38.

The results clearly show that the substitution of new pairs for those incorrectly responded to on any given trial retards acquisition under the particular conditions of Experiment I.

EXPERIMENT II

Subjects. The total of 80 *Ss*, volunteers from the introductory psychology courses, consisted of 43 men and 37 women. As in the previous experiment the *Ss* were assigned to the treatment-combinations in random replications. Each of 4 groups contained 20 *Ss*.

Procedure. The apparatus and stimulus-materials were the same as in Experiment

I. A single rate of presentation conforming to that reported by Rock was used throughout the present experiment.[3] During the learning phase, the slides were exposed for 3 sec. with a 5-sec. interval between successive slides. For the recall-test, the slides were exposed for 5 sec. with a 1-sec. interval between them. The method of replacement and rearrangement of the syllable-pairs was the same as before.

The main independent variable orthogonal to the control and substitution-conditions was the mode of *S*'s response. Half of the *S*s were instructed to spell aloud both the stimulus- and the response-members of the nonsense syllable-pairs while the remaining *S*s were instructed to pronounce both members of the pair. Thus, the four groups are identified as pronounce-control (*PC*); pronounce-substitution (*PS*); spell-control (*SC*); and spell-substitution (*SS*).

Results. Fig. 2 shows the trial-by-trial learning curves for each of these groups. The over-all level of performance under the conditions of Experiment II was better than that exhibited by the *S*s in Experiment I. Of the 20 *S*s in each group, the numbers that reached the criterion of one perfect recitation by the tenth trial were 18, 13, 13, and 7 for the *PC, PS, SC,* and *SS* groups respectively.

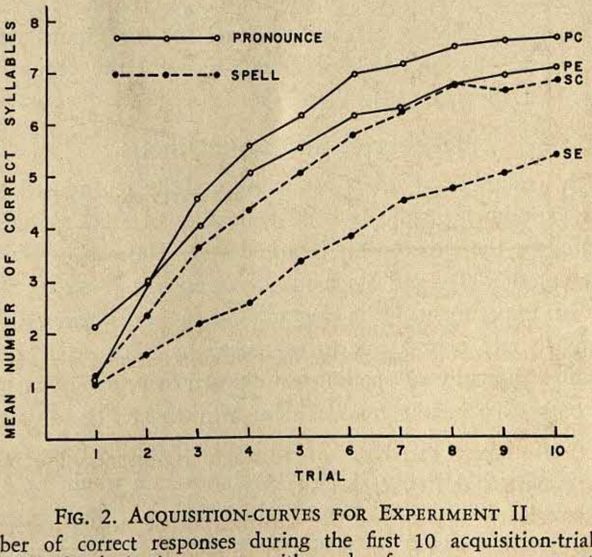


FIG. 2. ACQUISITION-CURVES FOR EXPERIMENT II

Mean number of correct responses during the first 10 acquisition-trials for the control and substitution-groups with mode of response as parameter.

A trials-to-criterion measure was obtained using the value 16 for *S*s who failed to learn by the sixteenth trial or exhausted the list of pairs before that trial. The means for the *PC, PS, SC,* and *SS* groups were 7.0, 8.4, 8.8, and 12.2 trials in that order. The tenth trial is the last trial for which complete data for all *S*s were available because two of the *S*s in the *SS* group had learned so few pairs by that trial that the list of additional pairs was exhausted. The total number of correct responses over the first 10 trials was therefore obtained for the analysis of variance. The means for the four groups were 57.8 for *PC,* 53.4 for *PS,* 49.1 for *SC,* and 35.6 for *SS.*

---

[3] Rock, *op. cit.*, 187.

The main effects of Mode of Response and Substitution-Conditions are significant at the 5% level. The interaction-term involving these variables fails to reach significance. Although the interaction is not significant, it is apparent that the difference between means for the control and substitution-conditions for the groups that spelled is very much greater than the difference between these means for the groups that pronounced: 4.4 as opposed to 13.5. Separate *t*-tests of the effects of substitution versus control means indicate significance at the 5% level for the groups that spelled ($t = 2.54$; $df = 38$) and lack of significance for the groups that pronounced ($t = 1.58$; $df = 38$).

DISCUSSION AND CONCLUSIONS

Although a number of investigators successfully reproduced the results of Rock's experiments,[4] the lack of generality of Rock's conclusions is demonstrated by the experiments reported here. The two variables investigated, presentation-rate and *S*'s mode of response, served to define task-difficulty. In Experiment I the degradation in performance even at the slower rate—a rate comparable to that used by Rock—can be explained by the greater difficulty of spelling as opposed to pronouncing. Thus, the results of both experiments are consistent with the idea that in a more difficult task, Rock's procedure will not result in evidence for one-trial learning.

A further test of the effects of presentation-rate would be to have *S*s pronounce under conditions similar to Experiment I. Such an experiment has been reported.[5] The difference between the replacement and non-replacement groups at the fast rate was highly significant, whereas no significant difference was found at the slow rate. These results support the above argument.

---

[4] L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, this JOURNAL, 73, 1960, 22-40; Michael Wogan and R. H. Waters, The role of repetition in learning, this JOURNAL, 72, 1959, 612-613; B. B. Murdock, Jr., and A. J. Babick, The effect of repetition on the retention of individual words, this JOURNAL, 74, 1961, 596-601.

[5] G. R. Lockhead, A re-evaluation of evidence of one trial associative learning, IBM Res. Rep., 1960, RC-279, 1-8.

More important, perhaps, than the differences among these particular experimental outcomes is the problem of identifying the nature of associational processes carried out by *S*. The definition of trial, stimulus member, response member, correct and incorrect may be completely operational for *E* in the traditional verbal learning experiment, but not so for *S*. A more fruitful approach to understanding the nature of assocition may well be the simulation-methods which use computer programs as models for the behavior of individual *Ss*.[6] The technique demands attention to what *S* is doing in greater detail than do the well-established paradigms oriented as they are to the precise specification of *E*'s behavior.

---

[6] Allen Newell, J. C. Shaw, and H. A. Simon, Elements of a theory of human problem solving, *Psychol. Rev.*, 65, 1958, 151-166; E. A. Feigenbaum, An information processing theory of verbal learning, *Proc. Western Joint Computer Conference, 1961.*

INFLUENCE OF JUDGES' CHARACTERISTICS UPON THE ADAPTATION-LEVEL

By DOROTHY RETHLINGSHAFER and E. D. HINCKLEY,
University of Florida

Judgments made in psychophysical measurement have been found to be influenced by more than the immediate objective stimulus presented. Helson has systematically approached some of the problems presented by such evidence in postulating that there is a point (or area) on any subjective scale, called the adaptation-level (*AL*), which represents the pooled effects of all stimuli other than the stimulus being judged.[1] He reported that a good first approximation to the *AL* in vision was obtained by taking a weighted logarithmic mean of all stimuli presented, forming the background in the field, times a constant.[2]

The present study is the prediction of the adaptation-level for a series of judgments upon ages of men and women, as well as a series of judgments upon heights of men, by assuming the judges' own ages (or heights) as background. In other words, by substituting such values in Helson's equations, the *AL* is predicted.[3]

Groups of individuals, each group of a different average age, were asked to judge a series of ages from 6 to 108 yr. The judgments were made on a 9-point scale from "very, very young" to "very, very old" with the value of '5' being assigned to "neither old nor young." It was found that the different groups of judges did not give the same scale-values. The younger the judges, the higher were the values reported. It was, therefore, hypothesized that the differing observed medium points (neither old nor young) might be predicted by an equation which took into account the average ages of the groups of judges.[4] In another study 'short' judges were compared to 'tall' judges in their evaluation of heights of men.[5] A consistent shift in scale-values was found, varying according to the heights of the judges.

---

* Received July 21, 1960.

[1] Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

[2] Helson, *op. cit.*, 2.

[3] Helson, *op. cit.*, 3.

[4] We wish to acknowledge the aid of W. D. Bliss, C. H. Maag, and G. Rettie, Jr. in the collecting of these data.

[5] E. D. Hinckley, and Dorothy Rethlingshafer, Value judgments of heights of men by college students, *J. Psychol.*, 31, 1951, 257-262.

Problem. This study was undertaken to determine whether a good prediction of such adaptation-levels could be made when the assumption stated above was followed; namely, the assumption that the judges' own ages, or heights, were background factors.

Helson has given the equation for AL in the following logarithmic form:

$$\log AL = \log K + [(k\ \Sigma \log X_1/N + \log C)/k + 1], \qquad [1]$$

in which K and k are constants, X is the value of a stimulus in the series, N is the number of stimuli in the series, and C is the value of the comparison-stimulus.[6] This equation was used in predicting the *AL*s for our recorded judgments, except that we found that the K term could be dropped in the judgment of ages, but was necessary to a small extent in predicting the judgments of heights. The assumption is made that the age, or height, of the judges represents the comparison stimulus designated by C.

Subjects. Judgments of age were obtained from 480 men and women divided into six different groups, with each group representing a different age-level. The size and average age of each group are presented in Table I. A series of ages was orally presented to the different groups and the judgments were made by the *Ss*' writing the appropriate number from a 9-point scale which was always before them. The younger groups worked together; the two oldest groups (Groups III and III B) were seen individually and responded orally. The judgments of heights of men were made by 260 men divided into two groups of tall and short judges.

TABLE I

SUBJECTS IN JUDGING A SERIES OF AGES OF MEN AND WOMEN AND SERIES OF HEIGHTS OF MEN

Judged series	Group	Sex of judges	*N*	Av. age	*SD*	Range
	I	boys and girls	53	9.7 yr.	.688	8–11
Ages of men	II	men and women	102	23.4 yr.	3.87	18–35
	III	men	54	77.0 yr.	4.84	67–88
	I B	women	99	10.1 yr.	.54	9–11
Ages of women	II B	women	148	18.91 yr.	.83	18–20
	III B	women	24	70.91 yr.	6.92	54–89
				Av. height		
Heights of men	tall	men	148	73.34 in.		
	short	men	112	66.9 in.		

Method. If the hypothesis is correct that the differing average ages of the groups can be considered as approximating differing backgrounds in determining scale-values, then it should be possible to predict theoretical scale-values for the adaptation-level from Equation [1] by substituting in this equation the appropriate value for C. The appropriate value for C would be the average of the group whose adaptation-level was being predicted. Equations were therefore written for each age-group. The following illustrates the equation for the youngest group of judges, whose average age was 9.7 yr.: $\log AL = (6\ \Sigma \log X_1/N + \log 9.7)/(6 + 1)$. Similar equations were written for the other groups judging ages. Note that the constant *K* was not found necessary to obtain a good prediction for the *AL*s in respect to ages. An additive constant was necessary to better the fit of the theoretical *AL* with the observed *AL* in the judgments of heights of men by men judges. The equation used to predict the *AL* for the 'short' judges was as follows: $\log AL - \log 2 = (5\ \Sigma \log X_1/N + \log 66.9)/(5 + 1)$. A similar equation was written for the 'tall' judges with the only change being in the *C* value.[7]

---

[6] Helson, *op. cit.*, 3.

Results. The results are presented in Table II. Formulae useful in predicting *AL*s for vision and for lifted-weights proved useful in predicting *AL*s for judgments of ages and also for judgments of heights of men. For the comparison-stimulus there is substituted a quantitative expression of one of the characteristics of the judges, the one most relevant to the judgment-situation. Examination of the predicted *AL* and of the observed *AL* indicates a high degree of relationship between the two. The greatest divergence is with Group III B, the 24 women whose average age was 70.9 yr.

TABLE II

THEORETICAL AND OBSERVED ADAPTATION LEVELS

Group	Av.	Predicted *AL*	Observed *AL*
I	9.7 yr.	36.3	36
II	23.4 yr.	41.2	42
III	77.0 yr.	48.9	52
I B	10.1 yr.	36.6	39
II B	18.9 yr.	40.0	41
III B	70.9 yr.	48.3	52.5
	Av. height		
Tall	73.3 in.	68.3	68
Short	66.9 in.	67.3	67.5

The closeness of the theoretical *AL*s to the observed *AL*s, in judgments of age and height, supports the assumption that the characteristics of judges are background factors in determining scale-values assigned by them to a series of stimuli. This assumption is implicit in Helson's theory.

Certain aspects of this report deserve careful consideration. First, the comparison-stimulus represented the average age, or average height, of a group. The group varied in homogeneity both in respect to the characteristic under consideration, and also in respect to other background characteristics. It might be assumed that any decrease in such variability would improve the prediction. Secondly, it should be noted that the fairly large weight found to be necessary for the geometric mean of the series judged may simply indicate that the series used, in judgments of both age and height, represented all possible stimuli usually experienced in these areas. The series of both ages and heights was quite wide in extent and included the entire range commonly experienced of ages and heights, and some not so commonly experienced. The *S*s, no matter what are their own ages, may bring a fairly well established point of 'neutrality.' This established center might not be shifted too easily by the order of presentation nor by a change in the series judged. A similar statement could be made in regard to heights.

---

[7] Constants were empirically selected in order to obtain the prediction of the *AL*. (See Table II).

Summary. A formula developed by Helson in predicting adaptation-levels (*AL*s) in the judgment of lifted weights was used for predicting them in the judgments of a series of ages and of heights. It was possible to predict theoretical *AL*s that were close to the observed *AL*s. In the use of the formula with the judgment of ages, the average age of each group of *S*s was considered as a comparison-stimulus in predicting the *AL* for that group. With the two groups of tall and of short judges, it was also found that the average height of a group could be used as a comparison-stimulus.

EFFECTS OF RETINAL FIGURE-GROUND RATIOS ON TARGET-SIZE AND APPARENT SIZE OF AFTER-IMAGES

By ALAN M. HARTMAN, VA Hospital, Palo Alto, California

Recent investigations of the relationship between Emmert's law and the law of size-constancy indicate that these laws are not converse, although the findings do not exclude the possibility of a converse relationship between them.[1] Edwards demonstrated that changes in apparent size of the after-image (*AI*) were not reported when the subjects (*S*s) attempted to estimate the sizes of *AI*s in a cue-reduced situation.[2] His results would be expected if changes in apparent size of the *AI* has some converse relationship to size-constancy. Hartman verified Edwards' findings, but also found that the major cue to the perception of apparent size of the *AI* in the cue-reduced situation was the constancy of the ratio of retinal areas subtended by the target and ground even though the target was an after-effect and did not vary in terms of area.[3] Since the Gestalt explanation of size-constancy[4] and some empirical evidence indicate that size-constancy depends to a great extent upon the ratio of retinal areas subtended by the figure-ground relationships,[5] the following hypotheses were drawn:

Hypothesis 1. If the size of the object in space is equated to the size of the projected *AI* in terms of retinal angles predicted on the basis of Emmert's law, there should be no difference in estimates of size located at the projection-plane and the after-effect which is perceived and projected at the same plane.

Hypothesis 2. Changes in the ratios of the retinal areas subtended by target and background or *AI* and the projection-plane via variation of the

---

* Received for publication December 19, 1960.

[1] T. G. Crookes, The apparent size of after-images, this JOURNAL, 72, 1959, 547-553; A. H. Hastorf and J. L. Kennedy, Emmert's law and size-constancy, this JOURNAL, 70, 1957, 114-116; F. A. Young, Boring's interpretation of Emmert's law, this JOURNAL, 63, 1950, 277-280; Concerning Emmert's law, this JOURNAL, 64, 1951, 124-128.

[2] Ward Edwards, Apparent size of after-images under condition of reduction, this JOURNAL, 66, 1953, 449-455.

[3] A. M. Hartman, The apparent size of after-images in delusional and non-delusional schizophrenics, this JOURNAL, 75, 1962, 587-595.

[4] Kurt Koffa, *Principles of Gestalt Psychology*, 1935, 222.

[5] Harold Bishop, Effect of variation in distance cues on the perception of object size. Unpublished Doctoral thesis, Washington State University, 1958.

area of the projection-plane alone would equally affect the size-estimates of the target and the *AI*. That is, as the ratios of the retinal areas are increased, estimates of the size of the target would be expected to incerase.

Hypothesis 3. As long as the ratio of retinal areas subtended by figure and ground remains constant, *S* would be unable to report changes in size or distance as the distance of the projection-plane varies.

It is the purpose of this study to check the verity of these hypotheses.

Procedure: (1) Subjects. The *Ss* were 28 staff members (psychology and medical) and their wives from the Palo Alto Veterans Administration Hospital. Each *S* was randomly assigned to one of four groups: Group A estimates the size of a target located 43 in. away; Group B, the apparent size of the *AI* projected 43 in.; Group C, the size of a target located 90 in. away; and Group D, the size of a projected *AI* located 90 in. away.

(2) Apparatus. The apparatus was a slightly altered version of the equipment used in a previous study and is described in detail elsewhere.[6] The basic units and alterations are listed as follows:

(a) Left-hand tunnel. The observational tunnel was an 8-ft. trapezohedral box, so constructed that the visual angle subtended by any perpendicular bisection remained constant.

(b) AI *induction-square.* A 1-in. square was cut into a mask which could be raised or lowered, in and out of the apparatus, 26 in. from *S*. The square was illuminated from behind by a 150-w. flood-lamp for 10-sec. intervals.

(c) Right-hand tunnel. This tunnel was 3.5 × 115 in. and was located to the right and parallel with the left-hand tunnel.

(d) Moving comparison-square. This unit consisted of a 1-in. square of translucent glass mounted in front of a light-tight box. The square was illuminated from behind by a $\frac{1}{4}$-w. neon bulb. *S* could move the square from his position in front of the apparatus through a pulley-sytsem. The range movement was 14 to 104 in. from *S*'s position.

(e) Cue-reduction mask. The mask was located 26 in. from *S* and contained a 3 × 3 in. opening for the left-hand tunnel and a 3.5 × 3 in. opening for the right-hand tunnel.

Since the distance of the *AI* induction-square was invariant, the apparent size of the projected *AI* could be predicted and estimated from Emmert's law. Furthermore, the areas of the target-patches were so calculated that they matched the expected physical size of the *AI*. At the 90-in. projection-distance, the target-square was $3\frac{9}{16}$ in.; and at the 43-in. distance, it was $1\frac{11}{16}$ in. The target-patches were attached to the ground glass plates by electrical tape inserts. Background areas were constructed of cardboard with the appropriate area cut from the center. The ratios of target to area, 0.89, 0.47, and 0.33, were obtained by making the background areas or projection-planes $1\frac{15}{16} \times 2$ in.; $3\frac{10}{16} \times 3\frac{11}{16}$ in.; and $5\frac{7}{16} \times 5\frac{9}{16}$ in., respectively, for the 43-in. projection-distance. At the 90-in. distance, the areas were $4\frac{3}{16} \times 4\frac{3}{16}$ in.; $7\frac{10}{16} \times 7\frac{12}{16}$ in.; and $11\frac{10}{16} \times 11\frac{12}{16}$ in.

(3) Method. The distance and ratio-presentation were completely randomized for each *S*, who received four trials for each ratio; two at 43 in. and two at 90 in.

The *Ss* were trained to use the moving comparison-square and wherever an *AI*-

---

[6] Hartman, *op. cit.*, 587-595.

series was indicated, they were given experience in developing and in viewing the *AI*. Only after each *S* indicated that he was familiar with the apparatus and understood what was required of him, was the series of trials begun. For each *S*, the first two trials of the randomized series were disregarded and then repeated after the completion of the series; so in all, there were 14 estimates of size by each *S*.

A face-mask attached to the front of the apparatus aided the *Ss* in maintaining a constant head-position and insured that all size-estimates were made monocularly with the right eye.

Results. The assumption of homogeneity of variance was not disproved and the data were accordingly subjected to a mixed type analysis of variance. The independent treatment of the estimates of the apparent size of the *AI*, the estimates of the size of the target-object, and the viewing-distances were not significant. Hypothesis 1 was confirmed. There was no difference in perceived sizes of targets and *AIs* which were equated in terms of retinal angles predicted on the basis of Emmert's law and then located or projected at the same distance.

The results also confirmed Hypothesis 2 that as the ratios of retinal areas were increased, estimates of the size of targets or projected *AIs* would increase. The effects of the ratios of retinal areas subtended by the projection and target-objects were significant at the $\frac{1}{10}\%$ level, while the interaction of the treatment-effects was not significant.

Hypothesis 3 was also verified. As long as the ratio between the retinal areas subtended by figure and ground was constant, *S* was unable to report changes in size or distance. No *S* reported distance-variation, and there was no statistical difference in size-estimates due to distance.

The results of this study indicate that estimation of size of the targets (both real and apparent), in terms of minutes of arc subtended at the retina by comparison-square, varies directly as the ratio of the retinal areas of stimulation become larger. Since the size of the target was constant, these results are due alone, as we believe, to changes (increase or decrease) in the size of the projection-field.

Discussion. Whether the target is a retinal after-effect of stimulation or the direct effect of light reflected from a target, *S* is able to make an estimate of size. In either case, this study demonstrates that perception of change in size depends upon deviations in the ratio of retinal areas of stimulation. Conversely, if the ratio of retinal areas subtended by figure and ground—the effects of distance controlled—is not varied, *S* is unable to report size-changes in the *AI*, or size-changes for any other form of a target. We conclude, therefore, that estimation of size-constancy and apparent size of *AIs* are related phenomena inasmuch as they depend upon the same factors or supporting cues. This is not the same, however, as stating that Emmert's law and the law of size-constancy are converse phenomena.[7] Whether these phenomena are conversely related depends in a great degree upon (1) how Emmert's law is interpreted and measured; (2) the constancy of the ratio of figure to ground subtended at the retina; and (3) the effective control of other cues to size and distance.

The present study is not a direct test of the relationship between the law of size-constancy and Emmert's law, although the results suggest that size-constancy or the lack of it may be highly related to the ratio of the angles subtended at the retina by the figure and the ground.

Summary. To test three hypotheses derived from the assumption that apparent size of the *AI* is conversely related to the perception of objective size (size-constancy) if the ratio of target or *AI* area to background area is varied in a cue-reduction situation, 28 *S*s were randomly divided into the following groups. Group A estimated *AI* size, distance 43 in.; Group B estimated the objective size of the target, distance 43 in.; Group C estimated *AI* size, distance 90 in.; and Group D estimated the objective size, distance 90 in. Each Group estimated the size of two targets, apparent or otherwise, at two distances for the following ratios of figure and ground: (1) 0.89; (2) 0.47; and (3) 0.33. The results were as follows: (1) Estimates of size were independent of the type of target; (2) Estimates of size were found to vary directly with the ratio of target and angles of the projection-plane (ground) subtended at the *S*s retina. Target-size always remained constant while the ground varied in area; and (3) Estimates of size of projected *AI*s and targets did not differ significantly with variations in distance.

We concluded, therefore, that perception of apparent sizes of *AI*s and size-estimates of a target located in space are dependent upon variations in the ratio of retinal areas of stimulation. They, consequently, may be considered to be related phenomena. Our results strongly suggest that Emmert's law and size-constancy are under the conditions of this study, converse phenomena.

---

[7] A. H. Hastorf and J. L. Kennedy, *op. cit.,* 114-116; Young, *op. cit.,* 277-280; Edwards, *op. cit.,* 449-455.

PERFORMANCE IN THE RUNWAY AS A FUNCTION OF STIMULUS-DIFFERENTIATION

By Eli Saltz, Robert N. Whitman, and Coleman Paul,
Wayne State University

Saltz, working in the area of verbal learning, has suggested that gradually rising learning curves do not indicate increasing strength of S–R association, but represent differentiation of the relevant stimuli and responses.[1] Once differentiation has occurred, association progresses rapidly. Against this point of view is the frequent finding that even in very simple learning situations such as running down a straight alley to a goal, where differentiation is presumably not a factor, as many as 80–120 trials are often necessary to attain asymptotic speeds.

The purpose of the present study is to investigate whether differentiation is a variable in performance in a runway. One obvious differentiation that can be conceived of as necessary derives from the fact that different responses are required in the alley vs. the goal-box. In the alley, the animal runs; in the goal-box it stops running and eats. While several factors might impede such differentiation between the behaviors in the alley and the goal-box (*e.g.* internal stimuli which are the same in both places), the present study is concerned with a simple, manipulable variable: differences in color. If the alley and goal-box are the same color (as is usually the case in studies using the runway), *S*s should have difficulty in differentiation, the response of stopping will compete with the response of running in alley, and the *S*s will run more slowly than if alley and goal-box were different colors.

The authors' position is obviously related to that taken by Lawrence and Miller in analyzing the results of several studies by Finger.[2] Finger reported that increased

---

* Received for publication August 7, 1961. The present study is a part of a program of research, supported by a grant from the National Science Foundation, concerned with the influence of similarity and differentiation of human learning.

[1] Eli Saltz, Response pretraining: Availability or differentiation?, *J. exp. Psychol.*, 62, 1961, 583-587.

[2] D. H. Lawrence and N. E. Miller, A positive relationship between reinforcement and resistance to extinction produced by removing a source of confusion from a technique that had produced opposite results, *J. exp. Psychol.*, 37, 1947, 494-509; F. W. Finger, The effect of varying conditions of reinforcement upon a simple running response, *ibid.*, 30, 1942, 53-68; Retention and subsequent extinction of a simple running response following varying amounts of reinforcement, *ibid.*, 31, 1942, 120-133.

number of reinforced trials led to faster extinction, contrary to Hullian prediction. Finger had, however, used a runway in which the start-box of one trial was the goal-box of the preceding trial. Lawrence and Miller reasoned that Finger's results were due to an increasing tendency for *S*s to stay near the start-box (since this is where they had just been reinforced) with increased reinforcements. In their own study, Lawrence and Miller used distinctively different start- and goal-boxes and found that increased number of reinforcements increased resistance to extinction. Unfortunately, the results presented by Lawrence and Miller did not conclusively show that the relationship between number of reinforcements and rate of extinction is a function of differentiation, since they did not have a control group in which goal-box and start-box were similar.[3]

The position taken by Amsel and Spence appears to be contrary to that of Lawrence and Miller and of the present writers.[4] Both Amsel and Spence assume that similarity between goal-box and alley in a runway will lead to greater generalization of conditioned anticipatory goal-responses to cues in the alley. Further, they assume that such anticipatory responses have motivational properties (K) that multiply existing habit-strengths of running in the alley ($K \times H$). Consequently, increased simularity between goal-box and alley would be expected to increase running speed. This prediction is opposite to that arising from the writers' position on differentiation. Amsel cites an experiment of Eninger as indirect evidence for his assumption of generalized anticipatory goal-responses.[5] Eninger, however, measured errors in a T-maze rather than speed down a runway; and Eninger manipulated similarity between goal-box and the arm of the T at the choice-point. This situation does not meet the conditions of the assumptions made by either Amsel or the present writers.

Method. The *S*s were 20 female albino rats of the Sprague-Dawley strain. Age at time of the experiment was 100–120 days. Five *S*s were randomly assigned to each of the four conditions.

A 2 × 2 factorial design was employed. The groups were (a) Black-Black: start-box plus alley black, goal-box black. (b) Black-White: start-box plus alley black, goal-box white. (c) White-White: start-box plus alley white, goal-box white. (d) White-Black: start-box plus alley white, goal-box black.

The runway used was 7 in. high, 4 in. wide, composed of a starting-box 7 in. long, an alley 4 ft. long, and a goal-box 7 in. long. The runway was of wood with a mesh wire-top. A rubber-flap hung from the top of the runway separating alley from goal-box so that *S* could not see the goal-box from the alley. The flap was the same color as the alley.

Two start-boxes, two alleys, and two goal-boxes were constructed, one painted white, the other black, so that the runway for any particular group could be assembled from the appropriately colored segments.

---

[3] That this may have been a serious lack of control is suggested by recent results (*e.g.* David Birch, J. R. Ison, and S. E. Sperling, Reversal learning under single stimulus presentation, *J. exp. Psychol.*, 60, 1960, 36-40) which indicate that rate of extinction is *not* monotonically related to number of reinforcements.

[4] Abram Amsel, The role of frustrative nonreward in noncontinuous reward situations, *Psychol. Bull.*, 55, 1958, 102-119.

[5] M. U. Eninger, The role of generalized approach and avoidance tendencies in brightness discrimination, *J. comp. physiol. Psychol.*, 46, 1953, 398-402.

The *Ss* were placed on a 23½-hr. schedule of food deprivation 4 days prior to the beginning of the experiment, and maintained on this schedule for the course of the experiment.

Measurement of running-time began when *S* was placed in the start-box and ended when *S* entered the goal-box. A dish containing one pellet of food weighing 45 mg. was at the far end of the goal-box, and *S* was permitted approximately 5 sec. to consume the pellet. *S* was then transferred to a cardboard-carton for approximately 10 sec. (during which time *E* recorded *S*'s running-time and added a new pellet to the dish in the goal-box). *S* was then replaced in the start-box for the next trial. Each *S* received 10 trials a day for 5 days. After each day's runs, *S* was placed in an individual cage for ½ hr. and permitted to eat *ad libidum*. Water was continually available between testing periods.

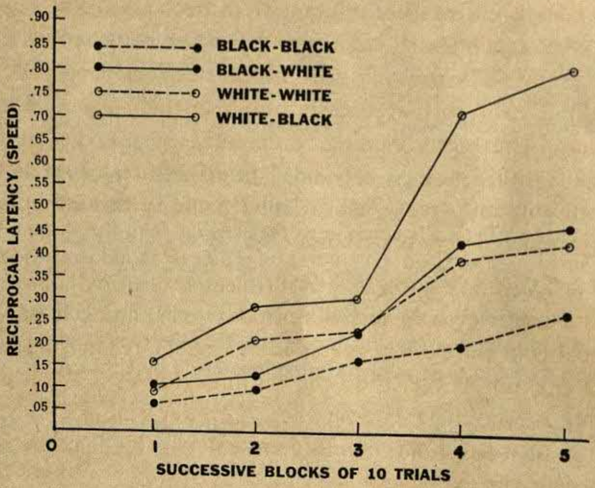


FIG. 1. MEAN RECIPROCAL LATENCY (SPEED) AS A FUNCTION OF DIFFERENTIATION BETWEEN COLOR OF ALLEY AND COLOR OF GOAL-BOX

The Black-Black and Black-White conditions were run simultaneously. After the 5 days of training for these groups, the White-White and White-Black conditions were run simultaneously.

Results. The results support the hypothesis that the traditional procedure for the runway, in which alley and goal-box are the same color, creates a problem in differentiation for the *Ss*. As can be seen in Fig. 1, the learning (as measured by running speed) of the Black-White group is greater than that of the Black-Black group; the learning of the White-Black group is greater than that of the White-White group. A 2 × 2 analysis of variance over the summed speeds for each group yields an *F* of 8.17 with 1 and 16 *df* for this comparison; this is significant at well beyond the 5%

level. In addition, the analysis indicates that the animals ran faster down a white alley than a black. This tendency is also significant at beyond the 5% level ($F = 8.30$, $1/16$ df).

Discussion. Voeks reports data which suggest that, even in classical eyelid-conditioning, problems of differentiation prevent rapid acquisition.[6] Voeks found that elimination of uncontrolled, irrevelant stimulation often produced one-trial jumps from zero to 100% frequency of CR. In the framework of the authors' position on differentiation, Voeks reduced the necessity for *S*s to differentiate relevant from irrelevant cues. These results are clearly consistent with the results and position of the present paper.

The present results, together with those of Voeks, suggest two conclusions. (1) Even in such apparently simple learning as occurs in the runway and in classical conditioning, differentiation may be an important variable. (2) If all factors impeding differentiation could be removed from the runway and from classical conditioning, learning curves might reach an asymptote in very few trials.

Summary. The purpose of this experiment was to determine if differentiation is a factor even in such simple tasks as running down a straight alley to a goal-box. Since different responses are required in the goal-box and in the alley (*i.e.* run in alley, stop and eat in goal-box) it was reasoned that running speeds would be slower when the alley and goal-box were the same color than when they were different colors. This hypothesis was sustained at beyond the 5% level.

---

[6] V. W. Voeks, Acquisition of S–R connections: A test of Hull's and Guthrie's theories, *J. exp. Psychol.*, 47, 1954, 137-147.

APPARATUS

A MACHINE CAPABLE OF BEING CONDITIONED

By JOSEPH R. CAUTELA, Boston College, and HARUTUNE MIKAELIAN, Brandeis University

Under the rubric of Cybernetics, Weiner includes such areas of study as the electrical engineering theory of the transmission of messages, the study of messages as a means of controlling machinery and society, and the development of computing machines and other such automata.[1] More specifically, it is the purpose of this paper to discuss the development of a machine that can be controlled and also the implications and heuristic value of such a machine.

The construction of the first fully electronic calculator at the University of Pennsylvania in 1945, and the later development of other calculators paved the way for the solving of problems heretofore too time-consuming or too difficult to solve. Another outgrowth of the development of such machines has been the increased speculation concerning the similarities between man and machine. One viewpoint is that many fruitful hypotheses can be generated about the structure and behavior of man because sometimes the behavior of machines can be shown to be similar to that of man, and that similar behavior may be determined by similar mechanism. The use of machine models in the investigation of man has one very distinct advantage over the use of lower animals; the construction and operation of the machine is known to a much greater degree than that of an animal. Also much more manipulation is possible with the machine.

Most of the speculation has developed out of the use of the digital electronic computer; the main assumption is that the covert behavior (thinking) of the human involved in mathematical computation is similar to the process used by the machine in the same computations.

More recently there has been some speculation and investigation of automata whose problem-solving behavior is not essentially of the mathematical nature but concerned with other types of adaptive behavior. One type of such automata is Ashby's "Homeostat" which simulates the human organism in seeking and maintaining homeostatic equilibrium.[2] Another such machine is the maze-solving machine developed by Shannon.[3] The kind of machine that is the concern of this paper is the type of conditioned reflex machine developed by Sluckin, who tells of a machine that was conditioned to respond to a whistle when paired to a light source.[4] A description of the machine was not given. Earlier attempts to develop learning machines that exhibit associative learning were of the one-trial learning type, hence no statistical data concerning the course of acquisition and extinction were presented.[5] We believed, therefore, that it would be worthwhile to build a machine that can provide data. Besides many other uses of building such a machine, we think it will make a contribution by providing statistical data, which have, as far as we know, never been reported by any investigator.

---

[1] Norbert Wiener, *The Human Use of Human Beings,* 1954, 15-17.

[2] R. W. Ashby, *Design for a Brain,* 1960, 100.

[3] E. A. Shannon, Presentation of a maze-solving machine, in H. Foerster (ed.) *Cybernetics,* 1954, 173-180.

[4] W. Sluckin, *Minds and Machines,* 1960, 64.

It must be borne in mind that the conditioned reflex-machine differs from the maze-solving and chess-playing machines developed in this country. These machines are problem-solving, whereas the conditioned reflex-machine behaves in a manner corresponding to the Pavlovian conditioning paradigm of straight classical conditioning.[6] If a neutral stimulus (which never alone elicits Response A) is paired with a stimulus that always elicits Response A and, after a number of pairings, the heretofore neutral stimulus elicits Response A, conditioning is said to have taken place.[7] In the typical Pavlovian conditioning paradigm, the bell (neutral stimulus—*NS*) is paid with food (unconditioned stimulus—*US*). After a number of trials the animal salivates (conditioned response—*CR*) at the sound of the bell (conditioned stimulus—*CS*) as well as salivating (unconditioned response—*UR*) at the presentation of the food. Extinction of the response takes place if the *CS* is not followed by the *US* for a number of trials. Spontaneous recovery occurs if the animal is allowed to rest awhile, after extinction, and then gives the *CR* when the *CS* is presented again.

THE MACHINE

A machine that fulfills the above conditions is described here. It (see Figs 1 and 2) is mounted on three wheels. The *US* is an activator—an extension with a button at (9)—which is connected to the apparatus and makes the machine move on its wheels (*UR*) when the button is pressed. The neutral stimulus is a beam of light focused on the front of the apparatus. If the light is paired with the *US* for a number of times, it (now the *CS*) will make the machine move (*CR*) before the presentation of the *US*.

After a number of presentations of the light without the *US*, the *CR* is not given and the machine is extinguished. If the machine is allowed to 'rest,' for a few minutes (3 or 4), and the light is again presented, the machine will move (*CR*). In other words, spontaneous recovery takes place.

An examination of Table I reveals that the number of trials to acquisition (the first time the *CR* equals the speed of the *UR*) and the number of trials to extinction (no perceptible movement) varies from session to session as it does in organisms.

---

[5] G. K. Bennett and L. B. Ward, A model of the synthesis of conditioned reflex, this JOURNAL, 45, 1933, 339-343; J. M. Stephens, A mechanical explanation of the law of effect, this JOURNAL, 41, 1929, 422-431; Albert Walton, Conditioning illustrated by an automatic mechanical device, this JOURNAL, 42, 1930, 110-111.

[6] I. P. Pavlov, *Lectures on Conditioned Reflexes*, G. V. Anrep, trans. 1927, 47-75.

[7] E. R. Hilgard, *Theories of Learning*, 1956, 50.

The machine does not reach acquisition at the different training sessions in the same number of trials because of certain variabilities. Conditioning of the machine is determined by a potentiometer whose shaft rotates at every impulse, thus decreasing resistance to current flow to the motor. This rotation, however, is not segmental, *i.e.* if the shaft receives a strong push, it rotates more than if it receives a weak push. The more impulses (*US*) the greater over-all extent of shaft-rotation. The extent of shaft-rotation at each impulse is, however, the critical determinant for the number of impulses required for conditioning. There are several variables which in

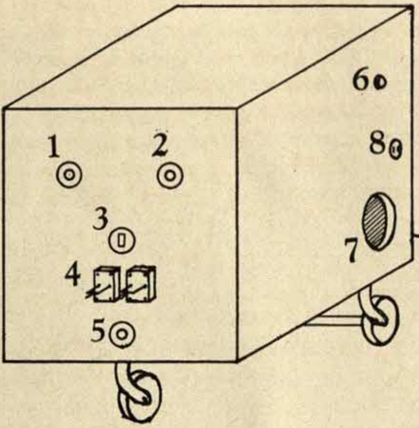


FIG. 1. DIAGRAM OF THE MACHINE

(1) Photosensitive control; (2) Spontaneous recovery delay; (3) Photosensitive element; (4) Switches for battery (AC power); (5) Photosensitive switch; (6) Pilot light indicating dormant condition; (7) Window to check extent of conditioning; (8) AC input; (9) Activator.

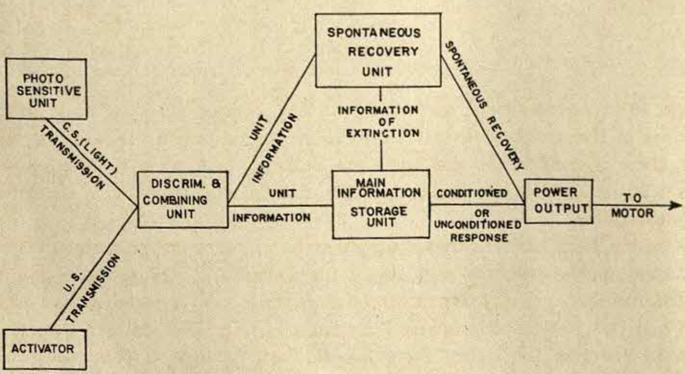


PHOTO SENSITIVE UNIT
C.S. (LIGHT) TRANSMISSION
DISCRIM. & COMBINING UNIT
U.S. TRANSMISSION
ACTIVATOR
UNIT INFORMATION
UNIT INFORMATION
SPONTANEOUS RECOVERY UNIT
INFORMATION OF EXTINCTION
MAIN INFORMATION STORAGE UNIT
CONDITIONED OR UNCONDITIONED RESPONSE
SPONTANEOUS RECOVERY
POWER OUTPUT
TO MOTOR

FIG. 2. BLOCK DIAGRAM OF THE MACHINE

turn determine this. First, the power of the battery, which is continually decreasing; the solenoid which transmits its translational movement into rotation of the shaft is pulled in by varying strengths. Secondly, the magnetic characteristics of the solenoid, the degree to which it holds its magnetism, until the impulse arrives is a critical factor, the solenoid pulls in strongly when initially the impulses reach it relatively frequently. After the fifth or sixth impulse, however, there is a rise in the temperature of the solenoid which once more changes the characteristics of the magnetic core and thus its pull. Again, the temperature-rise is a function of the impulse-frequency and cooling characteristics of the solenoid. The combination of the above mentioned factors accounts for most of the variability.

The machine operates on a series of relays and memory-systems. There are two stages built in the machine. The first stage accomplishes the initial conditioning, and the second accomplishes spontaneous recovery. The machine does not show retention from one training session to another because it is so built that, upon completion of the spontaneous recovery-cycle, it resets itself to its original condition. That is, there is built in the machine a mechanism which, upon completion of spontaneous recovery, erases the 'traces' of conditioning and resets it to its initial preconditioned state.

TABLE I

THE NUMBER OF TRIALS TO ACQUISITION AND TO EXTINCTION FOR 10 TRAINING SESSIONS

Training sessions	Trials to acquisition	Trials to extinction
1	6	65
2	7	56
3	10	49
4	6	66
5	8	48
6	6	55
7	8	50
8	9	57
9	13	60
10	8	54
	Mean 8.1	Mean 56.0
	Median 8.0	Median 55.6

Pressing the push button closes a simple series circuit with the motor thus making the machine move. Synchronizing this with a flash of light permits the photo-controlled memory-system to transform this impulse into a certain value of resistance. When this resistance is lowered beyond a certain value, enough current passes through to energize the motor every time a flash of light is presented; then the machine has been conditioned. At this stage, however, if the light is presented alone, the process is reversed in the memory-system. Now a certain value of resistance is added instead of being subtracted, so that after a while the resistance increases to the point where no current passes through when a flash of light is presented (extinction).

The value of the resistance that is added in each step when light alone is presented is one-fifth that of the resistance which is subtracted when the light and push button and light to condition the machine, then it would take 25 presentations of light (conditioned stimulus) for extinction. When complete extinction has taken place, the delaying relay is energized which closes after about 5 min., and transfers the operation to the second stage for spontaneous recovery. This consists of a stepping relay, which is energized by a flash of light connecting a set of batteries in series to the motor. When the last poles of the stepping relay is reached, the operation is transferred to the first stage, and the cycle can be repeated.

The machine at present exhibits only classical conditioning, reinforcement, extinction, and spontaneous recovery. The addition of external inhibition was deemed necessary, and the unit for accomplishing that has just been finished. As yet, however, it has not been added to the machine, as other additions also are being planned. When added, external inhibition will take place if, during the process of conditioning, a loud voice or intense vibrations are presented. The unit is made up of a sound-actuated relay which, when energized, will temporarily shut off the conditioning system.

Another addition that is being planned is the substitution of a sound-stimulus for the push button as the unconditional stimulus. When this is done, the machine will always respond if a certain frequency of sound is presented. Stimulus-generalization could also be added, but this would require bulky instrumentation, and since the machine operates purely by batteries, it was decided to leave out stimulus-generalization, but instead add another, and far more interesting, characteristic, maze-solving.

As yet the design for maze-solving has not been completed. There are several problems to be surmounted; paramount among them is the bulk of instrumentation, as it is intended to keep the machine as small and light as possible and yet still be powered by batteries.

Comparison of machine and animal. As far as the experimenter is concerned there is no essential difference between conditioning an animal and conditioning a machine in the sense that *E* presents stimuli, observes responses, and attempts to establish a relationship between them. In both types of *S*s (machine and animal), the receptor apparatus receives particular stimulation from an external environment and this stimulation is transformed to control entropy (measure of disorganization) through feedback. In both cases, responses are made on the basis of past experiences and the *S*s exhibit variability in their responses (*i.e.* the variability in responses for the different series of acquisition and extinction in our machine).

For our purposes, the value of the machine is related more to the differences between it and organisms than with the similarities.) One important difference involved in conditioning a machine is that much more is known about the nature of the principles operating the machine, and also of its construction than in the case of animals. Another important difference is that, in general, we have a more precise knowledge of the external stimuli influencing the behavior of the machine. The machines are not so sensitive to these and similar variables that influence organisms. Related to this difference is the factor of greater control of the past history of the machine in comparison to that of the animal. Another interesting difference is that there is a greater possibility of altering the machine to investigate different variables and to alter the conditioning process.

Value of the machine. The differences between conditioning an animal and a machine point to some interesting advantages of the machine. Spence has emphasized that in other sciences, analysis and investigation of simple situations under artificially controlled conditions quite remote from reality have played a decisive part in the successful development of comprehensive systems of knowledge.[8] He then makes a plea for the study of simpler types of learning with the use of lower animals. The development of the conditioning machine takes us one step further in control and simplicity and one step closer to the frictionless rolling balls and the perfect vacuum of the physicist.

With the notable exception of Ashyby's speculations,[9] the machine models of the human brain have so far been developed largely on the basis of the digital computer[10] The use of the conditioning machine as a model would seem more appropriate since organisms are more likely to form associations between sensory impressions and responses than performing mathematical manipulations. We know what is built into the machine to enable it to be conditioned; perhaps by analogy we can infer something about the structure and function of organisms which make these organisms capable of being conditioned. Of course, we must distinguish between a machine that can only imitate the overt human behavior from one that also works internally on the same principles as the brain.

It is also possible that with the adding of more and more conditioning phenomena, such as generalization, discrimination, and external inhibition, the machine will elicit all the known conditioning behavior of organisms, perhaps new laws can be discovered concerning the behavior of organisms by using only machines. This possibility is not as remote as it at first seems, for it is quite likely that if any conditioning procedure is described adequately in operational terms, then it can be built into a machine (this of course would include instrumental conditioning). MacKay, in discussing the possibility of imitating human behavior with the use of an artificial mechanism, puts it another way: "As soon as the required behavior has been specified by an adequate statistical test, a simple logical process can convert the test into a description of at least one possible mechanism that will meet it."[11] Also on this point, Ashby states that the essential difference between the brain and any machine yet made is that the brain makes extensive use of principles hitherto made little used in machines.[12] He further states that by the use of these principles (and he claims to know what these are) a machine's behavior may be made as adaptive as we please.

---

[8] K. W. Spence, *Behavior Theory and Conditioning,* 1956, 23.

[9] Ashby, *op. cit.,* 1-12.

[10] A. M. Turing, Can a machine think?, in J. R. Newman (ed.) *The World of Mathematics,* 4, 1956, 2101.

Once most of human conditioning behavior is built into the machine, hypotheses concerning the function of the brain and its relationship between external events and observable behavior can be formulated and tested on human *S*s. The use of the machine model might lead us to look for facts which might not otherwise have been considered with the use of another model or conceptual scheme. Then, too, there is always the chance that perhaps some mistake in wiring or other defect in the apparatus will lead to some behavior or some hypotheses previously unnoticed or unformulated.

Finally, the machine can be used very effectively, in the classroom, for demonstrating conditioning phenomena. The machine has the advantage of being able to elicit more conditioning behavior in a shorter period of time than with the use of an animal. The machine could be so constructed as to exhibit acquisition, extinction, spontaneous recovery, generalization, and external inhibition, all in one class period. The greater control over the machine will enable students to observe the conditioning process more closely than would be possible with the use of animals. For some small psychology departments that do not have animals, the machine would be of especial value. For other departments, the machines could be used in conjunction with animals; the machine could be used at first to demonstrate in a more detailed manner the conditioning process then a similar demonstration could follow with a use of an animal.

---

[11] D. M. Mackay, On comparing brain with machines, *Amer. Scientist,* 42, 1954, 261-268, esp. 261.

[12] Ashby, *op. cit.,* 1.

CONTACT OCCLUDERS FOR RESTRICTING VISION IN THE RAT

By JOHN R. SCHUCK, Bowling Green State University, and
LOUIS A. COPPOLA, Mount Kisco, New York

This note describes a technique of restricting vision in the rat, similar to that reported for the monkey.[1] After many trials with different sizes and thicknesses of plastic occluders, one was finally developed which could be inserted under the eyelids of the rat without injury to its eyes. The occluder, 0.01 in. in thickness, covers about half of the corneal surface. In many respects it is similar to the corneal lens used by man.

To test the occluders, they were inserted in the eyes of rats on 15 successive days and were permitted to remain there for 23 hours as previous testing had indicated that this was probably the maximum for safe use. Corneal clouding and vascularization occurred typically after periods of 24 to 72 hr. of continuous wear.

The *S*s were male hooded rats, about 180 days of age. They were housed individually. Feeding, watering, and general cage-maintenance were undertaken during the hourly period each day when the occluders were removed. At the end of the experimental period, every animal was given an ophthalmological examination; no abnormality nor injury was revealed in the eyes of any of the animals.

The method of inserting and removing the occluders was relatively simple and is illustrated in Figs. 1A and C. For insertion, the occluder was first moistened on both sides with a standard wetting solution for contact lenses and then placed on the tip of a small ear syringe, where it was held by suction. The eyelids of the experimental animal were gently drawn back from the cornea with the thumb and forefinger and the suction of the syringe released as the occluder was touched lightly to the central corneal surface.

Removal was accomplished with a small rubber cone mounted on a handle. The tip of the cone was inserted gently under the edge of the occluder at the point nearest to the external canthus of the eye and the occluder lifted off. In Fig. 1, translucent white occluders are used to illustrate these procedures. The behavior of the rats gave no indication that either the

---

* This work was conducted at the Laboratory of Comparative and Physiological Psychology, Ohio State University, under Grant MF 9095 from the National Institute of Mental Health.

[1] Mortimer Mishkin, R. D. Gunkel, and H. E. Rosvold, Contact occluders: A method for restricting vision in animals, *Science,* 129, 1959, 1220-1221.

insertion or removal was painful in any way. After the first few sessions, all animals were docile during both procedures. With several days of practice, the two experimenters were able to insert occluders into both eyes of the 12 rats in a total time of from 10 to 12 min.

For molding the occluders, a steel ball bearing was used as the male die and a piece of brass tubing as the female die. The tubing was fastened in the chuck of an ordinary drill press and the ball bearing was mounted on the platform. Small squares of vinyl plastic sheeting were then heated over a flame until sufficiently pliable for molding.

One of the major problems during the final operations was that of so holding the occluder that it could be ground and polished. This was successfully accomplished by the use of a surgical aspirator, the occluder being held fast against the end of a small rubber tube by suction.

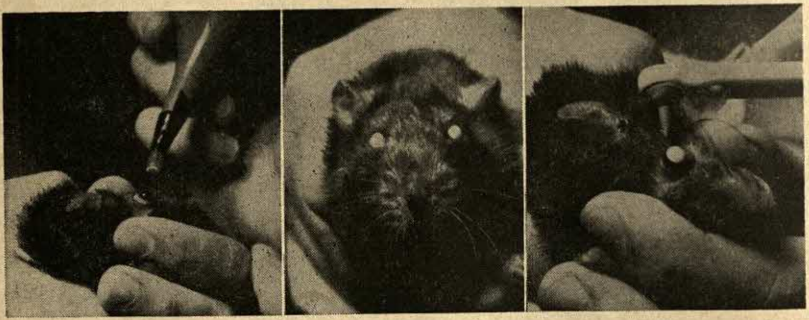


FIG. 1. INSERTION, NORMAL POSITION, AND REMOVAL OF OCCLUDERS

Finishing was done with a hand motor tool and accessories. Occluders were trimmed to the proper size with emery sanding disks. Ball cutters and grinding spindles were used to remove plastic burrs at the edges. Further smoothing and polishing were accomplished with felt cones impregnated with an acrylic plate polish and finally by felt buffing wheels.

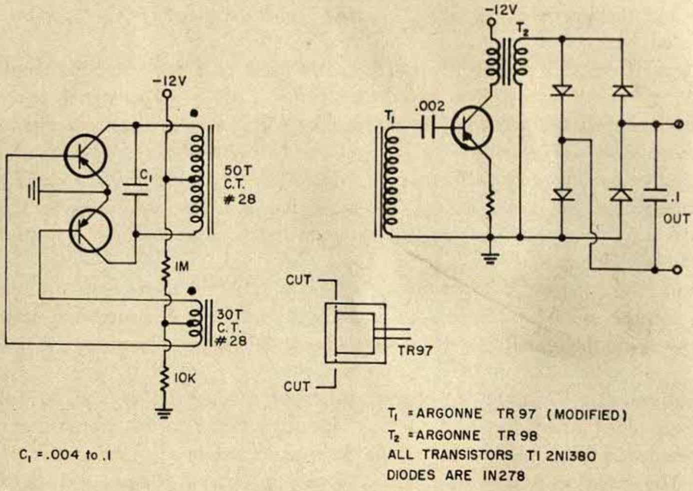
Occluders which were too large and consequently touched the apex of the cornea were found to cause clouding. To prevent this, the proper size was determined by placing a drop of a solution of fluorescein sodium on the concave surface of a transparent occluder, inserting the occluder in the eye of an experimental animal, and inspecting under ultra-violet light. Those portions touching the cornea appear dark blue or green; those not touching appear bright yellow.

In view of these results, contact occluders seem to provide a highly flexible method of manipulating the visual environment of the rat. "Blindness" can be induced without enucleation or other surgical procedures and normal vision can be readily restored. Clear occluders should provide reliable controls. Whether similar occluders can be developed for periods of continuous use longer than 23 hr. daily remains a problem for further investigation.

APPARATUS NOTES

APPARATUS FOR MEASURING RESTRICTED RANGES OF LINEAR MOVEMENT

The measurement of linear motion within restricted ranges such as that done in studies of body or limb movement has usually been based upon the mechanical transformation of the motion to a dial or a recorder.[1] One disadvantage in this type of measurement arises from the necessity for coupling *S* to the measuring instrument, thus restricting *S*'s free motion. The apparatus described below eliminates this disadvantage. Motion is converted to changes in AC or DC output-voltage picked up by a portable receiver from a fixed transmitter. The input to the receiver is then coupled to a meter or recording device. Within certain limits, changes in this output-voltage approximates a linear function of the motion produced and thus provides a direct measure of linear movement with no restriction on *S*'s free movement.



C_1 = .004 to .1

T_1 = ARGONNE TR 97 (MODIFIED)
T_2 = ARGONNE TR 98
ALL TRANSISTORS TI 2N1380
DIODES ARE IN278

FIG. 1. WIRING DIAGRAM OF PROXIMITY TRANSMITTER AND RECEIVER

The apparatus consists of two parts: a transmitter mounted in the proximity of *S* and a receiver attached to *S*. A schematic drawing of the circuit is shown in Fig. 1.

The transmitter radiates sinusoidal *RF* signals in a very low frequency-range, 600 cycles to 5 kilocycles. This is emitted from the coil parallel to its central axis and is roughly toroidal in configuration.

The transmitter is a multivibrator with a tickler winding providing base bias. The tickler coil (30 turns *CT*) must be phased properly to create oscillation. The transistors are alternately cut off and saturated to produce roughly a square wave in the collector load-coil. This is made sinusoidal by the parallel effect of *C*, which tunes the transmitter to the required frequency. With the values shown, transmitter frequency is approximately one kilocycle.

---

[1] A. S. Edwards, New apparatus for the measurement of bodily movement, *J. exp. Psychol.* 25, 1939, 125f.; Static atasiameter for head and hips, this JOURNAL, 54, 1941, 576.

The receiver in Fig. 1 uses an Argonne transistor input transformer (100K to 1K) modified by removing one side of the laminations so that the remainder presents an *'E'* configuration. The low impedance leads are clipped off and the output checked at various distances from the loop of the transmitter. An appropriate value of tuning capacitor is connected across the high impedance windings to peak the output. The value of this capacitor will depend upon the operating frequency chosen, the actual inductance of the windings (variable from one transformer to another) and the interwinding capacitance. At higher frequencies this capacitor may be dispensed with.

The signal from this pickup coil is fed to the base of a one-transistor amplifier, the collector load of which is a driver transformer (AR98). The signal generated in the lower impedance secondary is rectified by a full-wave bridge and filtered by the capacitor shown. The entire receiver unit can be built within a ½ × ½ × 1 in. volume and encapsulated in an Insulex *(R)* dip. Three leads are brought out, a common ground, negative power in and, the DC signal out. The amplitude of this DC signal will vary from approximately 2 v. immediately adjacent to the transmitter loop, to 20 m.v. at 3 ft.

The transmitter loop may be wound on a form suitable for the operation intended. For foot lift, the coil is wound on a 12 × 24 × ½ in. form. For use as an indicator of body sway of the standing person, the same or larger perimeter may be used.

Applications. To illustrate the characteristics of this apparatus, the DC voltage output was determined for (1) various distances between the transmitter and receiver and (2) nine placements of the receiver in relation to the center of the antenna. The distances ranged from 0 to 100 cm. in units of 10 cm., and the placements were the center and the eight 45° positions around the center, one-half the distance between the center and outer edge of the antenna. The antenna was 54 × 24 cm., and the receiver was always positioned perpendicular to the plane of the antenna loop. Output voltage was measured with an oscilloscope in units of 20 m.v. For all locations the decrease in output voltage for distances between 10 and 40 cm. approximated a linear function of the distance with 1 cm. in movement equalling 28 m.v. change in output.

The approximation is most accurate when the range of movement is smallest as in the case of limb or finger movements; but it is also adequate for measurement of body sway, since the error of measurement introduced by conventional recording devices or scales on voltmeters with a 2-v. range is usually sufficiently great so as not to justify computed adjustments for specific distances of linear motion. When used for measurement of body sway this apparatus has the advantage of not requiring adjustments for differences in heights of the *S*s, providing the *S*s are so placed in a position that the receiver is perpendicular to loci near the center of the antenna, *i.e.* closer to the center than to the outer edge.

VA Hospital
Perry Point, Maryland

C. C. Brown
J. R. Whitman

A VISUAL RATING SCALE FOR PRINTS OF FINGER-SWEAT

An index of finger-sweat has within recent years been widely used in work dealing with physiological measures of activation.[1] This method recommends itself for reasons of ease and convenience, particularly since the introduction of finger clamps for the purpose of taking such measures.[2] In a recent study, a visual rating scale for prints of finger-sweat was used, representing a further simplification of the method.[3] This scale is the subject of the present paper.

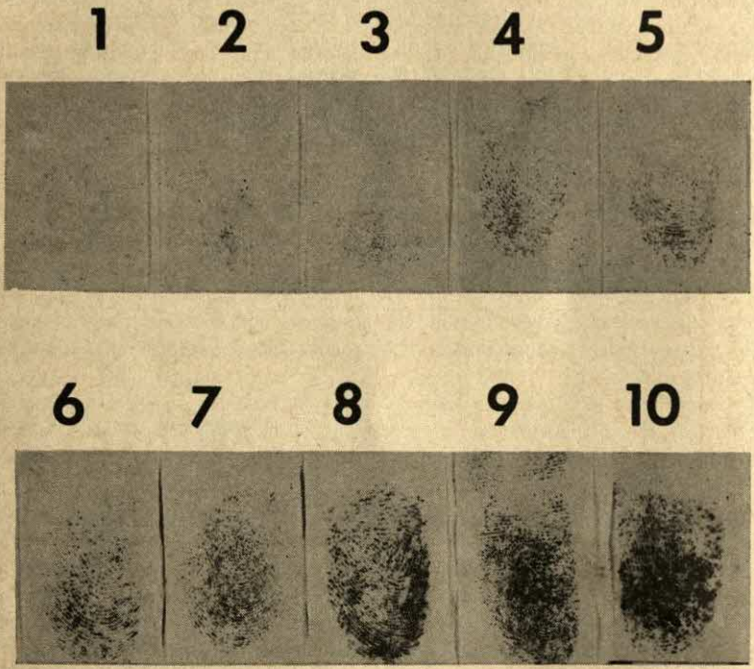


FIG. 1. GRADED SERIES OF PRINTS USED IN THE RATINGS

---

* From the Allan Memorial Institute, McGill University. This research was supported in part by the Department of National Health and Welfare, Defence Research Board (Canada), and the United States Public Health Service.

[1] J. J. Silverman and V. E. Powell, Studies on palmar sweating: I. *Amer. J. Med. Sci.,* 208, 1944, 297-305; Studies on palmar sweating: II *Psychosom. Med.,* 6, 1944, 243-249; Roy Gladstone, *An Investigation of Certain Variables Affecting the Results of a Group-Test of Palmar Sweating,* Final report, Sub. Contract No. HumRRo 650-005 with George Washington University, 1954; R. C. Wilcott, Silverman-Powell index of sweating vs. skin conductance and a humidity index of surface moisture, *J. comp. physiol. Psychol.,* 62, 1959, 33-36.

[2] D. A. Chambers, Rowena Pasternak, and H. F. Müller, A clamp for finger-sweat prints, *Perceptual and Motor Skills,* 11, 1960, 35-38.

[3] H. F. Müller "An exploratory investigation of 'clinical-type' physiological indicators in the measurement of arousal," Thesis for the Diploma in Psychiatry, McGill University, 1959.

The procedures for obtaining the sample of prints used in this study were the the same as those previously described by Chambers *et al.*[4] The samples included a wide range of prints differing in darkness. A selection of prints was made for a visual 10-pt. scale with subjectively equal darkness intervals, including the lightest (No. 1) and the darkest (No. 10) prints available. Densitometric values, obtained for the prints, were plotted against the visual rank-order. Photographic copies were then made of this visual scale.

A new sample of 100 prints was taken, and each print was classified according to its correspondence with the darkness of the scale-prints. For each print, the values obtained from the densitometric readings, the visual scale, and the photographic copies of the scale were compared. Test-retest and inter-judge reliabilities were computed.

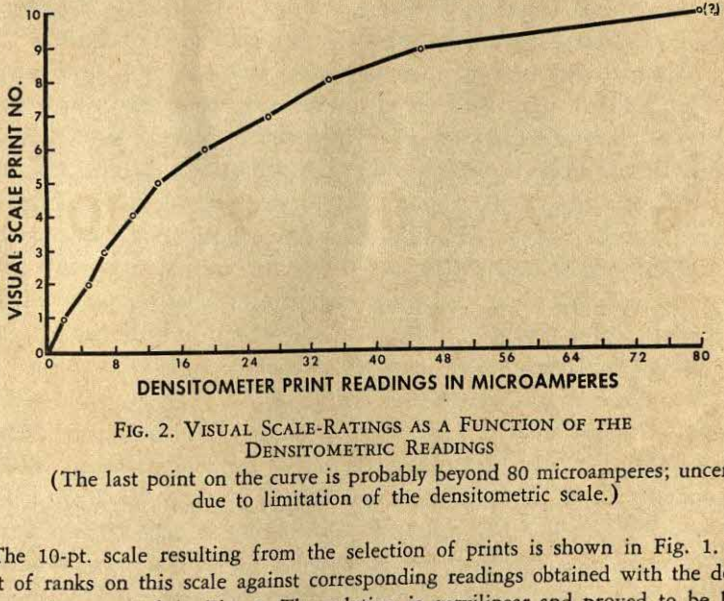


FIG. 2. VISUAL SCALE-RATINGS AS A FUNCTION OF THE DENSITOMETRIC READINGS

(The last point on the curve is probably beyond 80 microamperes; uncertain due to limitation of the densitometric scale.)

The 10-pt. scale resulting from the selection of prints is shown in Fig. 1. The plot of ranks on this scale against corresponding readings obtained with the densitometer is presented in Fig. 2. The relation is curvilinear and proved to be logarithmic. This was evident from a plot of the logarithms of the densitometric readings against the corresponding visual ranks, which resulted in a straight line for the range between Ranks 2 and 9. Obviously it is a Weber-Fechner function.

The various analyses for estimating validity and reliability yielded high correlation coefficients. The correlations between the densitometric values and ratings on both the visual scale and the photographic copy of the visual scale were respectively $r = 0.97$ and $r = 0.92$. The ratings on the photographic copy and on the visual scale correlated $r = 0.97$, and the test-retest correlation for the photographic copy was $r = 0.99$. The coefficients of inter-judge reliability (two judges) for the densitometer, the visual scale, and the photographic copy of the visual scale were $r = 0.96$, $r = 0.97$, and $r = 0.95$ respectively. (The *N* of prints was 100 in every case.) In addition, a reliability rating was obtained for three judges using the photographic scale. An analysis of variance technique was applied to the ratings and yielded a value of $r = 0.93$ for the reliability of ratings for one rater and $r = 0.97$ for the reliability of ratings for three raters ($N = 100$).

---

[4] Chambers, Pasternak, and Müller, *op. cit.*, 36-37.

Discussion. The results of this study show that finger sweating may be measured as reliably by a visual scale as by a densitometer. Such a scale has obvious advantages: It requires little equipment and is considerably faster. In addition, as many photographic copies of the visual scale can be reproduced as desired. The information obtained with the visual scale is adequate for most testing situations. As may be judged from Fig. 2, quite fine discriminations may be made with the visual scale, particularly at its lower end. This is an advantage in practice, because the majority of ratings fall in the range of light prints.

McGill University

JOHN KENYON
HERBERT F. MÜLLER

NOTES AND DISCUSSIONS

FAMILIARITY OF LETTER-SEQUENCES AND TACHISTOSCOPIC IDENTIFICATION

An observer (*O*) can identify fewer letters of an unfamiliar sequence than of a familiar sequence when such sequences are presented tachistoscopically for exposures too brief for complete identification. The information *O* receives from an exposure, however, is approximately constant when redundancy of letter sequences is taken into account. One of the clearest demonstrations of this fact has been produced by Miller, Bruner, and Postman.[1] The purpose of this note is to suggest a slight modification in the calculation of informational content of their stimulus-material, which will be shown to lead to an improvement in the orderliness of their results.

Miller *et al.* constructed pseudo words (sequences of eight letters) at zero, first, second, and fourth orders of approximation to English according to the procedure described by Shannon.[2] These pseudo words were presented to *O*s tachistoscopically for durations ranging from 10 to 500 m.sec. The *O*s identified and reported as many letters as they could.

When computing the information Miller *et al.* used both Shannon's and their own estimates of the redundancy of such sequences and argued as follows.[3] Letters chosen at random are all equally likely and the redundancy of the zero-order sequence is 0%. Letters in the first-order approximation are selected with frequencies approximately equal to those found in printed English. These frequencies so vary from letter to letter that the average amount of information per letter is less by about 15%. The information in the first-order sequence is thus 85%. The estimate of the redundancy of the second-order approximation is based on the distribution of frequencies of letter pairs as they occur in printed English. This reduces the information to 71%. The estimate of the redundancy of the fourth order is 43%, and hence the relative amount of information per letter here is 57%.

Two different measures of *O*s' performances in identifying letters were used by Miller *et al.*: *a placement-score* and a *letter-score*. To test the hypothesis that the amount of *information* received by *O* per exposure is constant, both scores were corrected for the average amount of information per letter by multiplying the scores by appropriate values of relative information. If the hypothesis is true and the increase in the scores is proportional to the decrease in the relative amount of information in higher-order approximations, then such a correction should eliminate variance in *O*s' performances attributable to familiarity of letter sequences.

---

[1] G. A. Miller, J. S. Bruner, and Leo Postman, Familiarity of letter sequences and tachistoscopic identification, *J. gen. Psychol.*, 50, 1954, 129-139.

[2] C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.*, 27, 1948, 379-423, 623-656.

[3] Shannon, Prediction and entropy of printed English, *Bell Syst. Tech. J.*, 30, 1951, 50-64.

The corrected letter scores of the *O*s in Miller *et al.* are plotted here in Fig. 1.[4] Although the curves are closer together than would be comparable curves for untreated letter scores, there is still some variability left. There is too great compensation for the redundancy at the second and fourth order. This led Miller *et al.* to suggest that "the letter scores do not provide a fair test of the hypothesis."[5] Better support for the hypothesis, however, was provided by the corrected placement scores.

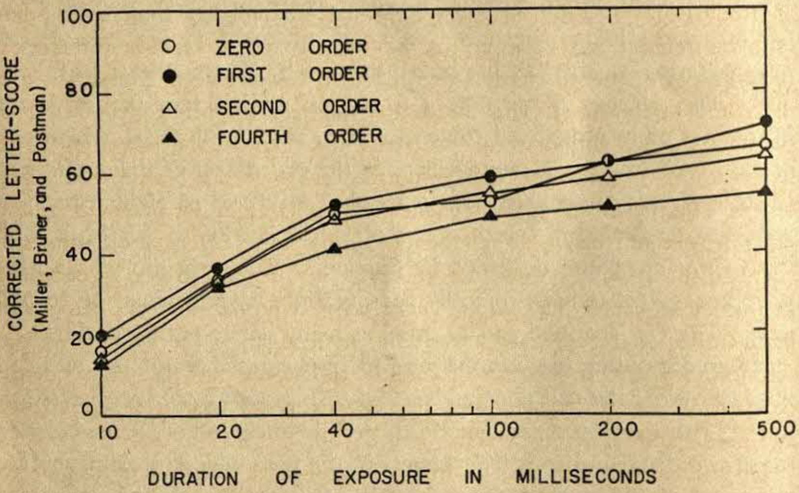


FIG. 1. CORRECTED LETTER-SCORES AS A FUNCTION OF DURATION OF EXPOSURE FOR SEQUENCES OF EIGHT LETTERS AT FOUR ORDERS OF APPROXIMATION TO ENGLISH

An examination of the corrected letter-scores in Table 3 of Miller, *et al.* makes clear that this residual variability is indeed orderly.[6] In that table all corrected letter scores for the fourth-order approximation are lower than those for the second order, and all corrected letter-scores for the second-order approximation are lower than those for the first order. Although this might logically be expected at relatively long durations of exposure, as Miller *et al.* point out, the reason for such an apparently systematic relation at short durations of exposure may lie in the underestimation of the relative amounts of information in the second- and fourth-order sequences.

---

[4] This graph is based on data in Table 3 in Miller, Bruner, and Postman, *op. cit.*, 137. An error in that table has been corrected here. The *letter-score* for the fourth order at 40 m.sec. exposure should be 41.6, not 46.6 as shown in the original table.

[5] Miller, Bruner, and Postman, *op. cit.*, 138.

[6] *Ibid.*, 137.

The estimates of redundancy in various sequences used by Miller *et al.* are based on the assumption of infinitely long sequences of letters. Such estimates are also correct for short messages of zero order and first-order approximations to English, but overestimate redundancy, and thus underestimate relative information, for higher orders. Take, for instance, a fourth-order sequence. Each letter in such a sequence is limited by the preceding three letters. This means that one can predict a given letter in such a sequence with maximal efficiency if he knows the preceding three letters. Suppose, however, that a fourth-order sequence contains only two letters. The first letter in the sequence of two letters is not preceded by any other letters and hence its redundancy and relative information value are equal to the average redundancy and relative information per letter in the first-order sequence. The second letter is preceded by only one letter and its informational content is, therefore, equivalent to the average of the second-order sequence. In such a case, of course, the fact that the sequence is a part of a longer fourth-order message is irrelevant as far as redundancy and predictability are concerned.

On the basis of this general line of argument it may be appropriate to revise the amounts of relative information in Miller *et al.*'s fourth-order stimulus-words. The first letter in the sequence is constrained by its individual frequency of occurrence in printed English, and contains, on the average, 85% relative information. The second letter is preceded by just one letter, its informational content being thus equivalent to the average of letters in the second-order sequence, viz. 71%. The third letter, by the same argument, can be estimated to contain, on the average, 63% relative information. For the remaining five letters the estimates based on infinitely long sequences, 57% relative information per letter, are quite appropriate. The average information per letter in a fourth-order sequence of eight letters, on the basis of the above reasoning, is thus estimated at 63%, instead of 57% as used by Miller *et al.* The same analysis applied to the second-order words yields an average relative information value of 73%. Zero order and first order are not affected.

Correcting the letter-scores in the Miller *et al.* experiment by multiplying them with these new values of relative information yields the functions plotted in Fig. 2.

Although there is still some variability left, the fit is decidedly better than in Fig. 1. Thus it seems that when more appropriate estimates of redundancy in higher-order approximations are used, Miller *et al.*'s corrected letter-scores also lend good support to their hypothesis. The amount of information received by an *O* familiar with the structure of the language is constant per tachistoscopic exposure.

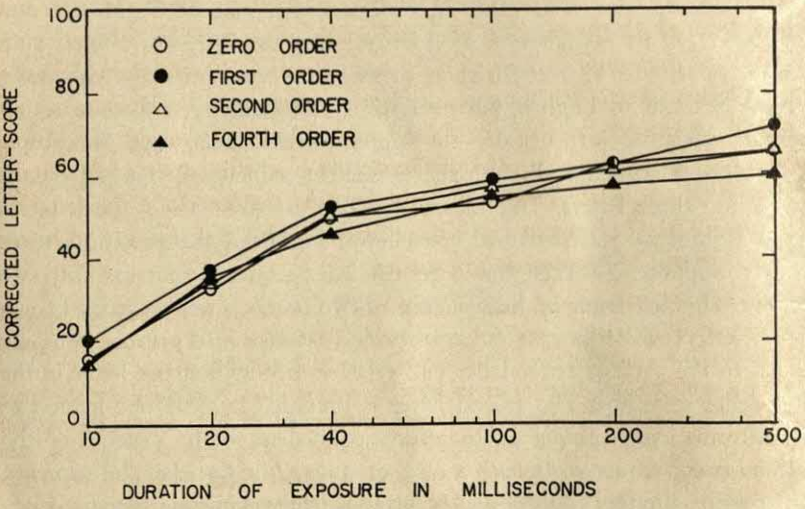


FIG. 2. DATA IN FIG. 1 REPLOTTED ON THE BASIS OF THE REVISED ANALYSIS

It should be mentioned that the revised estimates of the amount of relative information do not improve the orderliness of corrected placement-scores, but the fit for those scores is no worse than that for the corrected letter-scores. Furthermore, Miller *et al.* would probably agree that other variables, beside the information content of letter-sequences, are more likely to operate when *O*'s task involves correct identification of the position of letters than when the correctness of position is disregarded.

University of Toronto ENDEL TULVING

LOCALIZED NORMALIZATION OF TILTED LINES

Gibson, and Gibson and Radner reported that a line tilted by a small angle either from the vertical (defined as a plumb line) or from the hori-

---

* This research was supported by a grant from the National Science Foundation.

zontal (perpendicular to a plumb line) will appear less tilted after a short viewing period.[1] Gibson said that this effect "can occur in a localized part of an otherwise stable visual field."[2] He attributed it solely to a tendency of the tilted lines to approximate one of the main dimensions of space, either the vertical or the horizontal. On the other hand, Köhler and Wallach implied that all such localized tilt-normalizations might be interpreted as figural after-effects: apparent displacements resulting from the proximity of a test-line to the location of a previously inspected line.[3] Prentice and Beardslee presented data supporting normalization as a substantial part of the Gibson effect,[4] but their results were reinterpreted by Heinemann and Marill as special cases of the figural after-effect.[5] Obviously, if Gibson's localized normalization is to be demonstrated as an effect apart from a figural after-effect, exposure conditions must be used that preclude the latter interpretation. The present experiment attempts this task using a technique similar to that of Prentice and Beardslee.

Method. The experiment was conducted in a room entirely blacked-out with the exception of inspection- and test-lines and a fixation-point. Two roughly parallel lines were formed by two 12-in. 40-w. Lumiline bulbs operated at 15 v. and covered with opaque tape so as to leave luminous slits 12 in. long and ⅛ in. in width. The mounting of the bulbs allowed rotations of the lines around their center-points in the plane of the display, frontal to the *S*. The centers of rotation were 12 in. apart and the axis joining them could be made either horizontal or vertical by rotating the entire mounting. The fixation-point was midway along this axis. When either one of the lines was used for inspection with the axis horizontal, it was tilted 10° from the vertical either clockwise or counterclockwise; with axis vertical, the one line being used for inspection was tilted 10° off horizontal either clockwise or counterclockwise. In all there were eight possible settings of the single I-line.

During his session, each of 32 *Ss* had eight trials, separated by 3-min. rests, each using one of the eight settings of the I-line. The order of trials was so arranged that an I-line tilted off the vertical alternated with an I-line tilted off the horizontal. The 32 possible sequences of a vertical and horizontal inspection-line occurred an equal number of times throughout the entire experiment.

During each trial *S*, with head fixed by a biteboard, fixated the dimly luminous fixation-point at a distance of 8 ft. At the start of each trial, both the I-line and the second (non-inspected) line were switched on and *S* adjusted the comparison-line to appear parallel to the I-line. Two adjustments were made and the comparison line was then set at the mean of the two adjustments. At this point the I-line was switched on and *S* inspected it for 60 sec. The I-line remained on for an additional 3 sec. during which the comparison-line was also switched on. When both lines were visible *S*, as previously instructed, reported the direction (either clockwise or counterclockwise) in which the second line would need to be displaced in order to make it parallel to the I-line (now the test-line in the sense of a figural after-effect). He was forced to choose one of the two alternative directions. When *S* reported that the second line had to be displaced either toward the horizontal or the vertical from which the I-line had been tilted, the result favored the hypothesized local normalization.

---

[1] J. J. Gibson, Adaptation, after-effect, and contrast in the perception of tilted lines: II. Simultaneous contrast and the areal restriction of the after-effect, *J. exp. Psychol.*, 20, 1937, 553-569; J. J. Gibson and Minnie Radner, Adaptation, after-effect, and contrast in the perception of tilted lines: I. Quantitative studies, *ibid.*, 20, 1937, 453-467.

[2] Gibson, *op. cit.*, 564.

[3] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. phil. Soc.*, 88, 1944, 269-357.

[4] W. C. H. Prentice and D. C. Beardslee, Visual "normalization" near the vertical and horizontal, *J. exp. Psychol.*, 40, 1950, 355-364.

[5] E. G. Heinemann and Thomas Marill, Tilt adaptation and figural after-effects, *ibid.*, 48, 1954, 468-472.

TABLE I

DISTRIBUTION OF *S*s' RESPONSES

Responses in favor	0	1	2	3	4	5	6	7	8
No. *S*s	0	0	1	1	12	7	7	4	0

Results and conclusions. Each *S* could have given anywhere from zero to eight favorable responses. Table I lists the number of *S*s who gave each of the possible numbers of favorable responses out of eight. The mean appears to be shifted clearly in the favorable direction. Testing the results against a binomial distribution gives a Chi square of 21.8, which is significant beyond the 1% level.

The results are consistent with Gibson's claim of localized tilt-normalization and do not appear interpretable in terms of the figural after-effect. This conclusion is not inconsistent with the claim that a non-localized tilt-normalization also occurs.[6]

Brandeis University RICHARD HELD

THE ROLE OF REPETITION IN ASSOCIATIVE LEARNING

In view of the recent increase in interest in the role of repetition in associative learning,[1] it may be useful to draw attention to a paper by Miller and McGill in which they present a statistical description of "free-recall in verbal learning."[2] The method differs from the method of paired-associates in that on each trial of free-recall verbal learning a list of words is presented to the subject after which he attempts to write down all the words he can remember. In order to develop their theory the pertinent assumption made by Miller and McGill is that "the probability of recalling a word on any trial is completely determined by the number of times the word has been recalled on previous trials."[3] In other words, after the first presentation, repetition without recall produces no change in the probability of recall.

---

[6] J. H. Rich and R. B. Morant, A two factor approach to the study of tilt after-effects, Paper read at Eastern Psychol. Assn., New York, April, 1960.

[1] Irwin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193; L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, this JOURNAL, 73, 1960, 22-40; W. K. Estes, Learning theory and the new mental chemistry, *Psychol. Rev.*, 67, 1960, 207-223.

[2] G. A. Miller and W. J. McGill, A statistical description of verbal learning, *Psychometrika,* 17, 1952, 369-396.

In the form of the model which most adequately describes the data three parameters are used: (*a*) the initial probability of recall; (*b*) a parameter describing rate of learning or acquisition; and (*c*) a parameter which may be interpreted as an aspect of extinction or forgetting. The initial probability of recall is a constant as though the presentations before the first recall merely provide an opportunity for an association to be formed rather than providing small increments to such an association. The parameter describing rate of learning or acquisition is applied only on trials on which recall occurs; if recall does not occur, the probability of recall remains the same. The parameter which describes or takes account of a decremental process has the effect of permitting the description of learning which is less than perfect, *i.e.* where an asymptote less than unity is reached.

As can be seen from the foregoing description of the model, it requires the prediction that changing items that have not yet been recalled will have no effect upon rate of learning, a prediction consonant with Rock's finding using the method of paired-associates.[4] This feature of the model may have been a factor in the lack of experimentation and discussion excited by it. It may have been assumed by many learning theorists that a model which dismissed the reinforcing value of repetition might be technically adequate but still theoretically implausible. When Rock contrasted the same kind of findings with predictions from current theories of learning, he aroused considerable interest. The only extensive discussion of Miller and McGill's model of which the writer is aware is given by Bush and Mosteller who present a modified analysis of some of the same data.[5] These writers also ignore the implications for modern theories of learning and concentrate on the technical statistical description. One of their additional contributions is to point out that the model provides a method of measuring the assumption that the words were equally difficult for the subject to learn and that by this measure the words were of approximately equal difficulty.

University of Oregon JAMES H. STRAUGHAN

---

[3] Miller and McGill, *op. cit.*, 369.

[4] Rock, *op. cit.*, 186-193.

[5] R. R. Bush and Frederick Mosteller, *Stochastic Models for Learning*, 1955, 217-236.

A CRITICAL NOTE ON THE USE OF THE HUMAN CENTRIFUGE

Spinning a human being in a centrifuge has four main effects on his sense organs: (1) It displaces the direction of the gravitational force acting on his body; (2) It increases the force acting on his body; (3) It rotates him about his own body axis every cycle of the centrifuge; and (4) It produces a gradient both in the direction and the magnitude of the force acting on his body. If he faces in the direction of motion, his utricles will be unequally stimulated.

The centrifuge was used in all the studies reported in the psychological literature solely as a means of studying the effects on the perception of the vertical of a displacement of the direction of the force acting on the body. Clark and Greybiel did mention the possible effect of the increase in the force but did not present any results relevant to this factor.[1] The following argument shows that it is both an irrational and an inefficient method of studying the effects of a change in the direction of the force of gravity.

Gravitational and centrifugal forces are physically indistinguishable. Consider an enclosed room. There are two ways in which the direction of the force acting on the room may be altered. It may be rigidly attached to a centrifuge and so spun that the resultant of the centrifugal force and the earth's gravitational force makes an angle of α to the walls of the room, or it may be so tilted that the angle between the direction of the earth's gravity and the walls of the room is α. These two procedures are identical in their effect on any object or person in the room so far as the direction of force is concerned. It follows that an *S* blindfolded or in an enclosed room cannot distinguish between the experience of the displaced direction of force when the room is centrifuged, and the experience he would have if the room were merely tilted appropriately. The expense and trouble of making a centrifuge are unnecessary and, because of the other factors this method introduces, it gives results which are invalid as pure measures of the effect of a displaced direction of the force acting on the body.

There are purposes for which the use of a human centrifuge is ideal—for instance, the study of the effects of increased gravity on human performance. When used for this purpose the cubicle would be allowed to swing out so as to keep *S*'s body axis normal to the resultant direction of force. Rotating an *S* about his own body axis is a sensible way of studying the vestibular reflexes. Wapner, Werner, and Morant used this method to study the effect of rotation on the judgment of the vertical.[2]

---

[1] Brant Clark, and Ashton Graybiel, Visual perception of the horizontal following exposure to radial acceleration on a centrifuge, *J. comp. physiol. Psychol.*, 44, 1951, 525-534.

The fallacy of using the centrifuge to study the effects of a changed direction of force seems to go back to Mach.[3] He formulated the hypothesis that when an individual is subjected to the resultant of centrifugal and gravitational forces he senses the direction of the resultant mass acceleration and regards this as the vertical. Mach's hypothesis could be translated into the terms of the tilted room experience by saying that when an *S* is tilted in the dark he can still set a line to the gravitational vertical. The Aubert phenomenon reveals that this is only approximately true.[4] Mach was working in the nineteenth century and at that time not so much was known about the nature of gravitational or centrifugal forces, nor was much known about the sense organs involved in the perception of the vertical. Mach's hypothesis could be understood as stating that gravitational and centrifugal forces are equivalent in their action on the sense organs. Since Einstein, we know that there is no way of distinguishing a gravitational field from a centrifugal field, hence it cannot matter whether we produce a change in the direction of the force on the body by tilting *S* or by rotating him.

In recent times there have been two groups of experimenters working on the human centrifuge. Noble and Clark and Graybiel working at Pensacola had their subjects set a line to the vertical in an otherwise dark cubicle rotated on a centrifuge.[5] Their results fitted Mach's hypothesis fairly well; some of Noble's data departed from the hypothesis in the direction of the Aubert phenomenon. Their results would have had more significance had they simply tilted their *S*s.

Mann, Berthelot-Berry, and Dauterive, who were associated with the Pensacola unit, did in fact tilt their *S*s in the dark.[6] They evidently did not realize that the two experiments were materially equivalent; the one done at Tulane by Mann *et al.* was a better way of doing what was attempted at Pensacola.

---

[2] Seymour Wapner, Heinz Werner, and R. B. Morant, Experiments on sensory-tonic field theory of perception: III. Effect of body rotation on the visual perception of verticality, *J. exp. Psychol.*, 42, 1951, 351-357.

[3] Ernst Mach, Physikalische Versuche über den Gleichgewichtssin des Menschen, *Sitzungber. Math. Nat. Kl. Acad. Wiss. Wien.*, III Abt., 68, 1873, 124-140; *Grundlinien der Lehre von den Bewegungsemfindungen*, 1875, No. 4, 1-127.

[4] Hermann Aubert, Eine scheinbare bedeutende Drehung von Objekton bei Neigung des Kopfes, *Virchow's Arch.*, 20, 1861, 381.

[5] C. E. Noble, The perception of the vertical: III. The visual vertical as a function of centrifugal and gravitational forces, *J. exp. Psychol.*, 1949, 39, 839-850; Clark and Graybiel, *op. cit.*

[6] C. W. Mann, N. H. Berthelot-Berry, and H. J. Dauterive, The perception of the vertical: I. Visual and non-labyrinthine cues, *J. exp. Psychol.*, 39, 1949, 538-547.

Graybiel and Brown (1951) studied the delay in *S*'s adjustment of a luminous line after he was subjected to a displaced direction of force on the Pensacola centrifuge.[7] This question too would have been better studied by tilting *S*.

A second group working in New York did experiments both with a tilting room and with a centrifuge.[8] The use of these two techniques was defended by the following statement:

> It should be noted that modification of postural factors by simply tilting the body—a technique employed in previous studies—is basically different from their modification by altering the direction of the force that acts on the body. Tilting the body may at worst make it more difficult to detect the direction of the gravitational vector and to use it as a basis for determining the upright. The direction of the gravitational pull, of course, remains unchanged and continues to be available as a proper basis for judgment. Changing the direction of the force on the body, on the other hand, directly alters the gravitational standard for judgment."[9]

This argument is fallacious, the difference between the two techniques is only apparent to the experimenter; from the point of view of *S*, the two procedures are identical (side effects apart).

It is remarkable that this fallacy has persisted for 100 yr., forming the basis of a great deal of elaborate experimental work and finding its way uncritically into the textbooks.[10]

Careful thought does not reveal any advantages in using the centrifuge. The tilted room and tilted chair are less expensive and more easily and accurately adjustable, they do not vibrate as a centrifuge must and the results are not contaminated by side effects.

University of Durham
Durham, England

IAN P. HOWARD
WM. P. TEMPLETON

A PAIN-SCALE: A REPORT FROM A VICTIM OF TIC DOULOUREUX

The pain-scale in tic douloureux reported here was not prepared for publication. It was sent to me merely for my information. I was deeply interested as it described intensities and qualities of pain that were far beyond my personal experience. As I thought it of general interest and worthy of publication, I suggested that it be placed in the Notes and Discussions department of this JOURNAL. The author demurred at first: "I am reluctant," as he wrote me, "to publish it over my name as it might appear to others that I am seeking sympathy which, I assure you, I am not." Further correspondence, however, brought his agreement to publish it anonymously. Since every article needs an 'author' for purposes of indexing and reference, we decided to publish it over my name as an editorial Note. His report is as follows:

---

[7] Ashton Graybiel and R. H. Brown, The delay in visual reorientation following exposure to a change in direction of resultant force on a human centrifuge, *J. gen. Psychol.,* 45, 1951, 143-150.

[8] H. A. Witkin and S. E. Asch, Studies in space orientation: III. Perception of the upright in the absence of a visual field, *J. exp. Psychol.,* 38, 1948, 603-614.

[9] Witkin, Perception of the upright when the direction of the force acting on the body is changed, *J. exp. Psychol.,* 40, 1950, 93.

[10] S. H. Bartley, *Principles of Perception,* 1958, 329.

During the summer of 1952, the writer (who shall hereafter be referred to as *O*), had a severe attack of tic douloureux resulting from the extraction of an impacted, upper left, wisdom tooth. He was in considerable distress for a period of six weeks. An aftermath of the extraction was a partially numb lower lip covering an area slightly larger than a silver dollar, beginning at the median line of the chin and extending to the corner of the left lip. The feeling in this area was similar to the numbness of a foot which has 'gone to sleep' together with the prickly sensation which accompanies the circulation of the blood. When *O* became fatigued or overworked, these sensations increased to the point of distraction. They would, however, disappear completely when he was feeling well and was pleasantly occupied. On occasion the symptoms continued for some time, resulting in substantial discomfort. This could usually be relieved by taking three 5-gr. aspirin tablets.

On March 10, 1962, when *O* was under a heavy work schedule, he became aware of a substantial increase in the sensitivity of the lip-area. This was followed by an ache localized in the upper left second molar, which gave him so much pain that he consulted a dentist. An examination for caries and abscesses revealed that the tooth and its neighbors were sound and healthy.

By then the level of pain had increased to the point where *O* sought relief periodically with aspirin and demerol. The next morning following the dentist's examination, the pain had reached massive proportions. A physician was summoned who, after a few days, turned the case over to a neurosurgeon.

During the ensuing nine days, *O* was unable to concentrate for more than a few minutes at a time and, as a result, became quite disoriented in space and time. For him the world shrank to the side of his face and his cheek bone. He failed to attend to what he was saying or to what others were saying to him, and was insensitive to people and to his surroundings. During the following five days, the pain was excruciating despite all medication.

Initially, the attacks came at regular intervals of 1 hr. and lasted for approximately 20 min., but as the days passed the periodicity increased from 1 hr. to 3-, 6-, 12-, and 24-hr. intervals. Originally, the pain began in his teeth, spread through the entire left side of his face, and then lodged in his left eye and ear. Pain in the teeth was massive and of longest duration. The earache was the most excruciating but occurred less often and for shorter durations than the facial attacks. The pain in the eye was like that of a bad headache, with the sensation that the eye was being pushed from its socket. It was, however, several orders of magnitude less painful than the pain in the face and ear. Based on *O*'s introspections, a scale evaluating the degree of pain in tic douloureux was developed. As shown in Fig. 1, five levels of pain were identified.

Level 1. At the first level, *O* is just conscious of the irritation; for example, the incipient stages of a toothache or of a boil. He could normally allay the discomfort with an aspirin.

Level 2. At the second level, *O* is clearly conscious of the pain. It is intermittent and exhibits a rhythm with peaks occurring approximately 72 times per minute, suggesting a relation to the heart beat. An approximation of its wave-form is shown in Fig. 2. The small secondary peak gives the impression of a backwash or counter-swing of a pendulum. *O* could ignore the discomfort at this level by concentrating on a technical problem, planning an exercise, playing a card or chess game requiring a high degree of attention.

Level 3. At the third level, *O* can no longer ignore the sensations; severe pain rather than discomfort is now sensed. Pain is fairly constant for short durations with pulses occurring 35-40 times per minute. The experience is comparable to that accompanying the extraction of a tooth when the effect of the anesthetic wears off. *O* could, at this level, obtain some relief by placing a 5-gr. aspirin tablet between gum and cheek at the second molar and letting it dissolve, by using an infrared lamp, and by placing a heating pad on the side of his face.

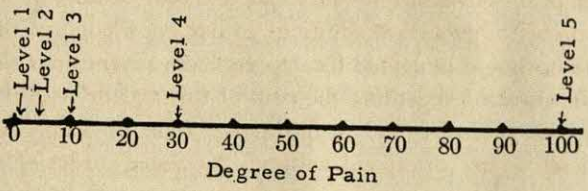


FIG. 1. SCALE OF PAIN IN TIC DOULOUREUX

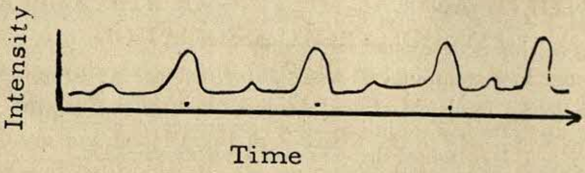


FIG. 2. TEMPORAL RHYTHM OF PAIN AT LEVEL 2

Level 4. The pain at Level 4 is continuous and severe. It is comparable to the continued pain that one experiences after hitting one's thumb with a hammer. Sublimation is no longer possible because one cannot concentrate except intermittently on anything except the pain. An aspirin held between gum and cheek and a heating pad provide token relief, as does inhaling Tryolene. Phenobarbital tablets and dylanthem permit *O* to doze fitfully but he builds up a substantial sleep deficit.

Level 5. At the fifth level, *O* loses contact with the world about him. The pain

> is so intense that his entire attention is concentrated upon it. He does not answer questions nor later remember that they had been asked. He paces the floor, apparently unseeing, as he frequently collides with the furniture and the walls. He can neither sit nor lie, but seems compelled to walk, perhaps an unconscious attempt to escape. At the height of the malady, pain of Level 5 occurs every 30 min., 20 of which are spent in pacing the floor and 10 in lying down exhausted. The attack builds up to maximal intensity in approximately 15 sec. and then declines as rapidly to Level 4. The pain in his teeth is comparable to that when a dentist's drill hits a nerve. This pain is combined with intense pains in the ear and massive dull pains in the eye. A graphic analogy of this stage is the fountain in a firework display: the continuous spray of fire bubbling from the fountain corresponds to Level 4; the fireballs and cannonades ejected intermittently from the fountain represents, Level 5.

O closed his report by cautioning the experimentalist against replicating this experience. No one, for the sake of gaining first-hand information, should attempt to self-induce it—he might succeed! Because of its severity, *O* is firm in his opinion that additions to our knowledge of the pain of tic douloureux should be left to the reports from adventitious cases. He is gathering information regarding the pain of this malady with the hope of constructing an intensive and a qualitative scale of pain. He will be pleased, therefore, to correspond with any interested reader of this Note; in particular with those suffering from tic douloureux in any of its forms. *O*'s name and address will gladly be given to those requesting it.

KARL M. DALLENBACH

SEVENTIETH ANNUAL MEETING OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

The American Psychological Association met for its seventieth Annual Convention in St. Louis, Missouri, August 30 through September 5, 1962. The sessions were held in the Chase-Park Plaza Hotel and in other nearby hotels. Only a fraction of the 9000 people attending the convention could find accommodations in the immediate neighborhood of the meetings themselves, however, so quite a number were lodged in other parts of the city. As has become customary, a small swarm of related activities hovers about the main convention itself. This year, for example, the Psychonomic Society held its third annual meeting at Washington University, August 30 through September 1. On the same days at the Park Plaza, the corridors seemed overflowing with the PIAP, Psychologists Interested in the Advancement of Psychotherapy. Actually, there were at least 15 formally organized groups not part of the APA who met in the 10 days before and during the APA meeting itself.

The focus of the APA meeting, as nearly as one can be found, comes on Sunday. In the morning, the three 1961 recipients of the APA Distinguished Scientific Contribution Awards, James J. Gibson, D. O. Hebb, and

Henry A. Murray, gave invited addresses. The afternoon session featured an address by Dr. L. C. Van Atta, Director of Research and Development Laboratories, Hughes Aircraft Co., on "Arms control: Human control." This was followed by the presidential address of Paul E. Meehl, on "Schizotaxia, schizotypy, and schizophrenia." Dr. Meehl suggested that we turn somewhat away from the recent trend toward a purely functional explanation of psychosis. The Gold Medal of the American Psychological Foundation was presented to Walter R. Miles. The 1962 Awards for Distinguished Scientific Contribution went to Jerome S. Bruner, William K. Estes, and Harry Helson. Only Dr. Helson was present to receive the award. An award of the American Psychological Foundation for outstanding science writing was given to Gay Gaer Luce for articles in the *Philadelphia Bulletin* and the *Saturday Evening Post*. Mrs. Luce is the wife of Prof. Duncan Luce of the University of Pennsylvania.

It is interesting to note that at some point in recent history this stopped being the Annual Meeting and became the Annual Convention. I leave it to the social anthropologist to give a full explanation of the changes in our mores that are reflected in this change in our language. Nearly 1100 stated papers were scheduled and there were over 1700 entries in the index of announced participants, about one-fifth of all who attended. The effect was not so much a town meeting of scientists but rather that of a gigantic state fair where psychological goods were exhibited. I shall not offer to chart where the prize cattle were to be found nor which were the grifter's wares.

The business of the Association comes to the Council of Representatives for final decisions. Important actions this year included final approval of a revised Code of Ethics, the decision to build a new headquarters building in Washington, an interim measure defining what psychologists could qualify for payments under medical insurance plans, and an expression of dissatisfaction with licensing laws that put state education officials in charge of Ph.D. programs in universities.

Dr. John G. Darley finishes a three-year term as Executive Officer of the Association and returns to his professorship at the University of Minnesota. Dr. Arthur Brayfield leaves the chairmanship of the Department of Psychology at Pennsylvania State University to take his place. Charles E. Osgood becomes President of the Association for 1962-63 and Quinn McNemar, President-Elect. The 1963 meeting will be in Philadelphia, starting on August 29 just three days after the close of the XVII International Congress in Washington.

Harvard University

E.B.N.

Ronald Aylmer Fisher: 1890–1962

Sir Ronald Aylmer Fisher, statistician and geneticist, died July 29, 1962, in Adelaide, Australia. He was one of a small handful of intensely creative Englishmen whose remarkable insights set the problems for a whole generation of mathematical statisticians. Scientists everywhere have exploited the methods of inductive reasoning that Fisher shaped and sharpened for them. His death will be keenly felt in many sciences and in many lands. Few sciences, however, have better cause to mourn his passing than psychology, for it was Fisher more than any other person—only L. L. Thurstone compares with him in this respect—who placed the powerful tools of mathematics at the disposal of even the humblest psychologist.

Of his many contributions to statistical theory, the most important to psychologists were analysis of variance and the design of multivariate experiments. So ubiquitous today is analysis of variance, so densely now do t-tests and F-ratios fill our journals, that it is difficult, especially for younger psychologists, to realize how recently we discovered Fisher's work. Only twenty-five years ago the design of experiments was still an intuitive art and there was almost no trace of small-sample statistics in our technical literature. In this brief span our whole conception of experimentation and data analysis has been revolutionized—a revolution all the more remarkable because we take it now so much for granted. Few psychologists have educated us as rapidly, or have influenced our work as pervasively, as did this fervent, clear-headed statistician.

R. A. Fisher was born February 17, 1890, in a suburb north of London. At Cambridge, he studied mathematics and in 1912 was a Wrangler, but his interest ran more to innovations than to axiomatics, more to applied problems than to pure mathematics. He began his study of statistics at the Mercantile and General Investment Company in 1913; his first statistical paper was published in *Biometrika* in 1915. After World War I, he accepted a post as statistician at the Rothamsted Agricultural Station and there he began to develop the methods that are known as analyses of variance. The identification of sums of squares with corresponding degrees of freedom was typical of his genius, so seminal, yet so deceptively simple.

The results of this work were first published in 1925 in *Statistical Methods for Research Workers,* but it was 15 years before they began to find their way into psychology. This was a difficult book and only a few copies of the first edition were printed, so it was not until the summer of 1931, when George W. Snedecor invited Fisher to visit the Statistical Laboratory at Iowa State College, that word of these remarkable new techniques began to spread widely in the United States. It was Snedecor who most actively introduced Fisher's ideas to Americans; it was E. F. Lindquist who most helpfully explained their significance for research in education and psychology.

Fisher's innovations in the design of agricultural experiments were first published in 1935 in *The Design of Experiments.* The randomized blocks now familiar to us as Latin squares were first laid out in the dirt of an experimental farm. Like so many of Fisher's insights, however this simple improvement had consequences extending far beyond the original problem that gave rise to it.

In 1933, Fisher was appointed to succeed Karl Pearson as Galton Professor in the University of London. There followed at London a decade of ardent, often violent, but consistently important work in mathmatical statistics by Fisher and his colleagues. Nevertheless, his long interest in the applications of statistical theory to genetics continued to develop, and in 1943 he accepted the appointment as Arthur Balfour Professor of Genetics at Cambridge University. Following his retirement from Cambridge in 1957, he took up residence at the University of Adelaide in Australia, where he continued to work until his death.

Harvard University GEORGE A. MILLER

ERRATUM

An omission in the last line on p. 463 of Dr. V. R. Carlson's article on "Underestimation in size-constancy judgments" (Vol. 75, 1962, pp. 462-465) alters the meaning of the sentence. Instead of ending "from objective matches," the sentence should end "from objectively correct matches."

K.M.D.

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Explorations in Transactional Psychology. Edited by FRANKLIN P. KILPATRICK. New York, New York University Press, 1961. Pp. xiv, 405. $7.50.

This collection of papers is a study of perception from that radically empiricist point of view, known as "transactionalism" which was the work of Adelbert Ames, Jr. and his colleagues, first at Dartmouth and later at Princeton. The book assembles fifteen previously published papers (1946-1956) and six new ones, of which a few present some genuinely new material. Although transactionalism is sometimes presented as having implications for psychology as a whole, virtually the entire range of evidence treated here comes from the study of space-perception. The phenomena discussed are the technically elegant, fascinatingly equivocal inventions and demonstrations of Ames and his collaborators, such as their studies of the perception of trapezoidal rooms, the rotating trapezoidal window, and the world seen through aniseikonic glasses.

Transactionalism is an elaboration of Helmholtz's doctrine of unconscious inference: the proximal stimulus is intrinsically equivocal, and the choices the organism makes among alternative perceptions are based on the purposes and assumptions he has developed in the course of his life history. An assumption is a "weighted average of past experience built up by action, checked by action, and modified by action" (p. 4). In spite of the blend of probablistic functionalism and developmental theory that transactionalism seems to promise, little or no attention is given to the work of authors like Bruner, Brunswik, Helson, Werner or Piaget. More important than these omissions are the failure to quantify a single probability and the failure to carry out a program of developmental research so that the claims of transactionalism could be tested more exactingly.

The central problem of the book is the assertion that a given proximal stimulus configuration may be produced by a large number of different objects, giving rise to the need for perceptual choices based on the individual's purposes and past experience. The transactional attack on stimulus theories of perception sets up something of a straw man: probably no psychologist, not even Gibson, claims that there is an unique relationship between a distal object and a single proximal stimulus configuration. Gibson's claim is, rather, that the many retinal stimulus-patterns produced by a single object as the organism moves in relation to it contain certain invariant relations, and that these invariants are the basis of perception: not specific configurations, but invariants that persist across many configurations. Unfortunately, the present work contains little of the close analysis in which such disputes might be clarified. One of the best chapters (Chapter 8 by Ittelson and Kilpatrick) does give a very clear account of the way in which a family of physical configurations can give rise to the same visual cues. This analysis is done separately for size, illumination, movement parallax, and binocular disparity, and could readily be done for other dimensions of stimulation. The authors recognize that it would be important for transactionalist theory to show that *combinations* of cues encountered in natural environments also produce equivocal stimulus patterns. But this proof is not given, although the assertion that "visual distance indications are *always* ambiguous" (p. 164) is the essential point of departure for the entire book.

The volume perpetuates a type of argument from demonstration that experimental psychologists need not necessarily admire. In the studies involving the rotating trapezoidal window, for example, it is argued that the peculiar phenomenon of physical rotation being perceived as oscillatory movement can only be explained as the outcome of assumptions made by individuals who expect to see and therefore do see rectangles whenever the retinal image belongs to a certain family of quadrilaterals. This argument can be tested by answering two questions. First, do individuals reared in a culture without rectangular forms see the rotating trapezoidal window differently from inhabitants of our rectangulophilic world? Secondly, are similar oscillatory effects obtained with shapes other than trapezoids? Both questions have been answered (the first by Allport and Pettigrew, the second by Pastore), and the experimental results are quite damaging to the claims of transactionalism. Neither these criticisms nor possible answers to them are mentioned in the book.

The book is full of unsupported appeals to an ecology containing probabilities or regularities about which the individual learns in the course of life history. Such "arguments from ecology" merit careful scrutiny. It is reasonable to suppose that a theory leaning heavily on variations in ecological probabilities should lead to experiments in which such probabilities are systematically manipulated. Neglecting this precaution leads to such contradictions as Ames' report on page 140 that an observer wearing aniseikonic lenses saw water as flowing uphill, followed on page 141 by an explanation of another observation on the grounds that "the observer knows from past experience . . . that the surfaces of bodies of water are always horizontal." From time to time the various authors indicate awareness of contradictions of this sort, but their work provides no basis for deciding whether the resolution of such cognitive conflicts actually involves probablistic weighting of different cues, or whether it can be fully explained in terms of immediately given stimulus factors.

Ames' report that in aniseikonic patients "alterations of anomalous binocular stimulus patterns cannot be brought about by experience and practice" (p. 145) offers at least some support for stimulus theories of perception and against transactionalism. Thus, in spite of the fact that the authors do not grapple decisively with the evidence contradicting their theory, be it said that the evidence is there for a careful reader to consider. The phenomenological descriptions of the world seen through aniseikonic lenses, of perception in distorted rooms, and of the appearances of the rotating trapezoidal window are given at some length and represent valuable contributions to perceptual research.

Some of the claims of perceptual empiricism are supported by experiments brought together in this volume. For example, Kilpatrick demonstrates clearly that visual exploration of events in a distorted room is sufficient to produce alterations in perceptual schemata; Wittreich shows that the particular personal relations existing between individuals influences the way in which they appear to each other in distorted rooms; Kilpatrick shows that the individual's assumptions about the material of which objects are made influence the perceived behavior of these objects in the field of a rotating trapezoidal window; Engel shows that perceptual conflicts in a binocular rivalry situation are resolved in favor of seeing faces in familiar orientations and not seeing faces in unfamiliar orientations; and Ittelson shows, in at least one of his experiments, that assumed size influences perceived distance. Experiments of this kind indicate that variations in past experience are *sufficient* to produce variations in perception, but they suffer from two limitations: first, the obdurate nativist can still argue that past experience is not *necessary* to produce an organized perceptual world; secondly, the experiments are not designed to shed much light on the particular way in which past experience operates.

One exception to the last point is Kilpatrick's attempt to compare the effectiveness of visual and motor explorations of a distorted room on the perception of such rooms. The experiment is well conceived and conducted, but the result was a surprise to Kilpatrick and other transactionalists who believe that perceptual assumptions depend on action. There was little or no difference between visual and motor exploration: both were effective. This result pretty well sums up the state of affairs: the volume as a whole leaves us not much wiser as to what *kind* of transaction is involved in perceptual development.

In his introduction, Kilpatrick insists on a transactionalist epistemology in which there is no reality independent of the participating organism. It seems to me that the same epistemology is as consistent with a "phylogenetic transactionalism" in which perceptual assumptions regarding important stimulus variables develop in the history of the species (nativism) as it is with an "ontogenetic transactionalism" in which perceptual assumptions develop only in the life history (empiricism). Consequently, transactionalism as represented by this volume is not a decisive theory of perceptual development but a limbo in which all sorts of theories can be invented and all sorts of experiments performed.

University of Colorado HOWARD L. GRUBER

Philosophical Theory and Psychological Fact: An Attempt at Synthesis. By CHARLES A. WALLRAFF. Tucson, University of Arizona Press, 1961. Pp. xi, 218. $6.00.

Now that philosophy and psychology have been working independently of each other for so many years, with inadequate communication and discussion between philosophers and psychologists, it is time to review what each field has to offer the other. Here we have a welcome attempt to do so. Wallraff is concerned with the possible contributions of the psychology of perception to the solution of the problems of epistemology, but is not concerned with the possible effects of philosophical investigations upon the psychology of perception. His book is therefore primarily for philosophers, but it is interesting to a psychologist to see what use he makes of psychological contributions as well as to learn something of current thought in epistemology, an area which we have difficulty separating from our own concerns. This review will not attempt to evaluate Wallraff's summaries of philosophical positions or his critical evaluations of these positions, except insofar as his criticisms depend upon his use of psychological theory and facts.

In stating his purpose, Wallraff says "I have been content to concentrate upon what seems to me to be a cardinal error of recent Anglo-American philosophy, with a view to showing that the coercive 'facts'—whether 'atomic,' 'brute' or simply 'hard'—which an entire generation of philosophers has been trained to accept at face value are actually hypothetical constructs, that the 'indubitables' of the armchair are not to be found in the laboratory, and psychological science has demonstrated once and for all that the cognitive enterprise is immensely more complex and far more engaging than the simple game of observation of givens and subsumption under laws which scientistic positivism has so frequently recommended" (p. viii). The author's title and his introduction would seem to say that he is going to substitute psychological or experimental "facts" for the armchair "facts" which he considers of doubtful accuracy. Yet he makes but little use of the kind of observation which the experimentalist thinks of as the raw facts of his discipline—the kind of observation which is repeatable and highly dependable. Instead, what Wallraff accepts from psychology are mainly conclusions and inferences at several removes from such laboratory observations of "facts." More important, few of these conclusions or generalizations are worthy of such confidence that we can take them as a basis of demolishing a philosophical theory. Some represent fairly general consensuses among present-day psychologists of perception, but even these generalizations lack the kind of solid foundation in experiment and observation which is characteristic of such an established theory as, say, the atomic theory.

Here are some of the ideas which Wallraff accepts and makes use of in his argument: (1) The Gestalt statement that classical sensations are not observed as prior to form, object, etc. and Katz's "modes of appearance" of colors as evidence for the conclusion. (2) Gibson's distinction between the visual "field" and the visual "world"—but without taking into account his accompanying objections to the "cue" or "sign" theory which is accepted by Wallraff. (3) The experimental studies of constancy by Kardos and by Brunswik are used in support of the notion of cue or interpretation theory. (4) According to the author, students of differential psychology and of cultural anthropology "showed that there is no such thing as 'human nature' *überhaupt*, and that consequently we have no reason to assume single *Weltanschauung* that comes naturally and easily to all peoples in all times and places" (p. 148). (5) The experiments of the Ames group of "transactionalists" are also taken to support the interpretation theory.

In all of these examples, there are experimental facts involved, in most cases quite extensive research is implied. Yet the facts are not allowed to speak for themselves and the interpretations which Wallraff gives are certainly not the unanimous conclusions of the psychologists in the field. Moreover, Wallraff has his own re-interpretations which will often come as a surprise to the author of the original conclusion. For example, he argues that there is support for the interpretationistic theory in the Getalt principles of organization, because the principles imply that sensory stimuli are actively organized by the perceptual process.

Wallraff's position seems summed up in the following statement: "Once the bewitchment of our intelligence imposed by the traditional mosaic of sensations has been overcome, we are in a position to recognize that knowledge, whether perceptual or conceptual, begins, not with the clear, distinct, and well articulated, but with the vague, amorphous, and enigmatic, and that knowing consists largely in imparting form and structure to this inchoate phantasmagoria. Above all, the world is thoroughly and incurably ambiguous" (p. 151). This reviewer confesses himself puzzled about the status of the "vague, amorphous, and enigmatic . . . inchoate phantasmagoria." Wallraff has accepted the Gestalt conclusion that what we perceive immediately is already organized, formed, meaningful, etc. At the same time he seems to say that what we perceive is formed *out of* something vague and meaningless. Isn't this the sensationistic doctrine which he is rejecting? True, he argues that the sensations of theory were supposed to be "*ex hypothesi* clear and perfectly determinate" (p. 154) while he finds sensations to constitute a "confused mass." The point is that *any* sensations which exist prior to organization are contrary to Gestalt theory—that portion of Gestalt theory which Wallraff seemingly accepts and quotes in support of his own position.

It appears that Wallraff has been less cautious and critical in his evaluation of the contributions of psychology than he is in analyzing the shortcomings of philosophical theories of epistemology. It may be that there is still some difficulty in communication which leads to misunderstanding. Possibly the philosopher may tend to depend too much upon the abstract verbal summaries of psychologists rather than becoming thoroughly familiar with the experimental methods and the facts derived from them.

Wallraff is to be congratulated for seeing the need to re-connect the lines of communication between epistemologists and psychologists. He has made a laudable attempt to summarize the results of the psychological studies of perception which are relevant to epistemological problems. Since psychological theories of perception are still highly tentative, it is natural that controversy will arise over the legitimacy of Wallraff's particular conclusions. It is hoped, however, that his study will arouse interest and further efforts in the same direction, and that we shall see aid frequently crossing the borders of the two fields in both directions.

T.A.R.

Conceptual Systems and Personality Organization. By O. J. HARVEY, DAVID E. HUNT, and HAROLD M. SCHRODER. New York, John Wiley & Sons, 1961. Pp. xii, 375. $7.50.

This book attempts to describe personality, its nature, development and pathology, in terms of specific modes of relatedness between individual and environment that the authors call conceptual systems. Concepts are systems of ordering and evaluating impinging stimuli; they range from the simplest judgments to the self, "The totality of one's concepts, of one's subject-object relationships, in their intertwined interdependencies" (p. 62). The particular conceptual system that is operating determines how an event will be experienced, and the response follows from this experience.

The structural properties of these modes of relating to objects are held to be independent of the nature of the objects. Conceptual systems vary along a number of dimensions, the most fundamental of which is abstractness-concreteness. They exhibit several additional structural properties, which produce variation in the degree of abstractness achieved: clarity-ambiguity, compartmentalization-interrelatedness, centrality-peripherality, and openness-closedness. Concepts have motivational properties—tendencies to approach or avoid their objects. Positive or negative affect follows conceptual confirmation or refutation; specific emotions are associated with threat to each major conceptual system.

Development is seen as a matter of attaining increasing abstractness of conceptual functioning; it may be arrested at any one of several stages of conceptual structure or transitional levels between stages. It proceeds by a process of "generating and integrating discrepant feedback or differentiations from the environment" (p. 85). The specific strivings whose successive conflicts and integrations at progressively higher levels constitute development are those of dependence and independence. Thus the progression is from symbiotic dependence through various stages to positive independence and interdependence; and at the same time a sequence from more concrete to more abstract functioning. Although both descriptions are plausible, the relation of these two modes of development is far from clear and is more assumed than spelled out by the authors. The same question arises in connection with the view that the various forms of psychopathology represent arrested development at different stages of abstract functioning. For example, schizophrenia is viewed as arrestation at System I (the most concrete) functioning. The thinking of schizophrenics has indeed been described as concrete; but it also shows a characteristic overabstractness and overgeneralization. Perhaps a more plausible hypothesis is that conceptual thinking develops, or exhibits difficulties, simultaneously in its handling of the abstract and the concrete.

System properties are assumed to apply alike to single concepts and to the total self system. At this point, the problem of organization seems to be glossed over. If the personality is made up of interrelated sub-systems, the problem of organization consists in part of finding concepts to describe their functional interrelations. This is not the interest of the present authors. Centrality is indeed an organizational concept, but it is used here mainly to refer to the openness of a conceptual system to external events, its strength, or its resistance to change—not to the relations of systems to one another. Compartmentalization is a dimension used with only somewhat more relevance to personality organization.

The specification of the nature of the organism, for the authors, "takes the form of dimensionalizing the structural properties of systems, which are then treated as the major internal factors" (p. 246). It may be doubted, however, whether a dimensional approach, however valuable for other purposes, can solve the problems of organization of personality, since the same dimension appears and functions so differently in the context of two different personalities. It seems necessary rather—just as in the case of perceptual organizations—to search for whole properties that describe the particular constellation of dimensions and other properties of a given individual. It would appear hazardous to assume at the outset that the relevant whole properties can be reduced to quantitative differences along specified dimensions.

Discussions of personality have not always given cognitive factors their due. Thus an approach based on conceptual functioning is very welcome. At the same time, one wonders whether justice can be done to affective factors in a presentation that centers on the concrete-abstract character of functioning. What this dimension means in relation to non-intellectual functioning is by no means clear. More generally, too much seems to be omitted from the self when it is viewed merely as a system of concepts; for example, even those unrealized potentialities out of which new concepts arise have no place in the system. The problems usually subsumed under unconscious processes are missing from this account. Motivational problems (except for dependence-independence strivings) are treated only schematically as the volitional (approach-avoidant) aspect of concepts.

In short, although this account makes no claim to completeness, it is hard to see how many important problems of personality can ever find their way into such a framework. On the other hand, the cognitive approach is a healthy, if perhaps exaggerated, corrective to many current orientations. The book offers categories that are refreshingly and stimulatingly different from ones conventionally used in the psychology of personality. Yet the discussion avoids being esoteric by showing relations with more prevalent conceptions (*e.g.* the relation of System I functioning to the authoritarian personality). The abstract language of the book suggests that it is written for a System IV audience. But the novelty of approach, as well as the importance of the issues raised, will reward the efforts even of those functioning more comfortably at a lower level of abstraction.

New School for Social Research MARY HENLE

Learning and Human Abilities: Educational Psychology. By HERBERT J. KLAUSMEIER. New York, Harper and Brothers, 1961. Pp. x, 562. $7.50.

This book represents a valiant attempt to break out of the traditional straitjacket that has rendered educational psychology over the past four decades a superficial, ill-digested, and often disjointed miscellany of learning theory, developmental psychology, educational and psychological measurement, mental hygiene, client-centered counseling, and child-centered education. Klausmeier, in this reviewer's opinion, correctly conceives of educational psychology as primarily concerned with the nature, conditions, outcomes, and evaluation of classroom learning, and organizes his material accordingly in terms of the principal variables impinging on these topics. The result is an impressively cohesive and well-integrated textbook which accomplishes the frequently attempted but rarely achieved aim of systematizing and unifying the field of educational psychology.

The emphasis on learning and cognitive abilities reflected in the title is, of course, in harmony with the *Zeitgeist.* Since the late 1930s, educational psychologists have been preoccupied with personality development, mental hygiene, sociometry, group dynamics, role playing, psychotherapy, and projective techniques. Despite the self-evident centrality of classroom learning and cognitive development for the psychological aspects of education, these areas were largely ignored, both theoretically and researchwise. Learning chapters in educational psychology texts were padded with wildly extrapolated findings from animal, maze, motor, rote, instrumental, and simple discrimination learning.

Unfortunately, however, Klausmeier is much less successful in making urgently needed substantive and theoretical innovations in particular areas of content than in redefining the proper scope and emphasis of educational psychology as a whole. Although one of the stated aims of the book is "to set forth a theory of classroom learning," there is no evidence anywhere of a consistent theoretical framework or point of view. Major theoretical issues in learning and personality theory are either evaded completely, or are treated very superficially—even from the standpoint of an introductory textbook. The book's strength lies, rather, in the bringing together of "research findings that deal with the efficiency of learning." Explanations, where required, are presented at a very proximate, cursory level, relying on sheer pragmatism or on a naïve eclecticism indifferent to the criterion of theoretical self-consistency.

Both the strength and the weakness of the author's approach are illustrated in his discussion of principles bearing on the learning of facts. Klausmeier considers such important variables as the organization of the learning task, the avoidance of initial errors, the reinforcement of correct responses, the immediate correction of errors, and the value of practice, review, and self-evaluation. He, however, makes no mention of perhaps the most significant variable of all; namely, the influence of the clarity, stability, and organization of related concepts and propositions in the learner's cognitive structure. This omission, of course, is reflective of a rote learning approach to meaningful verbal learning, which recognizes no theoretical distinction between rote and meaningful learning processes; and the research evidence cited in this section (pp. 173-182) is drawn almost entirely from studies of nonsense-syllable even though *most* facts taught in the classroom are relatable in a meaningful way to a previously learned body of general principles.

Greater stress is placed on cognitive development than in most other educational psychology texts, and the material is skillfully integrated with related aspects of cognitive functioning. Here again, however, it lacks theoretical sophistication and incisiveness. Criticism of Piaget's work, for example, is uncritically accepted without reference to the basic theoretical issues involved. Furthermore, no mention is made of Piaget's crucially important research in this field conducted since 1950. The outstanding chapters in the book are those concerned with "understanding and nurturing the range and variety of emergent abilities of children and youth." There are also excellent chapters on new instructional media and educational methods, and on the common statistical terms and procedures required for intelligent reading of the research literature and for the evaluation of classroom learning. Lastly, unlike many other texts in this field, the illustrations are carefully chosen for their ability to enhance understanding of the material rather than to provide relief, diversion, sentimental atmosphere, or an aura of scientific precision.

On the whole, the author has produced a sound, readable textbook which presents the relevant research findings in an honest, straightforward fashion. He does not ignore limiting factors or methodological qualifications in the research literature for purposes of enunciating wider and neater generalizations, or of advancing his own biases. In this reviewer's opinion, Klausmeier's book will easily secure an established place for itself among the better-known and respected texts in the field of educational psychology.

University of Illinois DAVID P. AUSUBEL

From Adolescent to Adult. By PERCIVAL M. SYMONDS with ARTHUR R. JENSEN. New York, Columbia University Press, 1961. Pp. xi, 413. $8.75.

Clinical research falls into an uneasy and unsatisfactory compromise of nomothetic and ideographic modes of thinking. As a scientific investigation of the fantasy life of young adults, or as a study of the predictive power of adolescent adjustment for adult status, this study is unconvincing. As a collection of detailed data on the fantasies and experiences of clinically unselected young people, the book is highly interesting. The authors are imaginative but plausible in their essentially subjective interpretations.

A follow-up of the senior author's earlier work on adolescent fantasy (*Adolescent Fantasy,* 1949), this effort reached 28 of the original 40 young people 13 yr. after the original study. It is based on a lengthy interview, responses to 20 of the original 42 thematic pictures, and a Rorschach protocol obtained from each subject. The present book is richly illustrated by excerpts from the protocols of these individuals, and includes a complete transcript of the interview with Jack, a case reported in detail in the original volume. Such materials will be of great interest to the clinically minded.

This reviewer would say that the most valuable contribution of the present volume is the survey of fanstasy themes in young adults' responses to thematic pictures. This survey suggests that young adults do not differ greatly from adolescents. There is an increase in themes of depression and wishful thinking; there is a decrease in themes of violent aggression, punishment for aggression and reform. Probably as a consequence of these changes, there seems also to be a decrease in themes of masculine identification. By a selective review of individual cases the authors satisfy themselves that in many instances the themes of adolescent fantasy work themselves out in adult behavior and adjustment.

The present volume attempts to study the transition from adolescence to adulthood. No new insights into this process appear, however. Adolescent girls develop a keen interest in boys before boys reciprocate this interest. Before marriage, some individuals have many friends of the opposite sex and others have very few of such friends. None of the young people had any formal preparation for marriage. Most of the young marrieds consider themselves happy and doing reasonably well; the Rorschach testifies that many of them are poorly adjusted and have inadequate sex-role identification. In few cases did their educational plans, stated in adolescence, work out as they expected. Vocational plans simply aren't made by adolescents. The typical adolescent falls into a job, his selection being primarily determined by the amount of money he can earn; he works out his vocational career from that point fortuitously. Symonds regrets that despite the many counseling and planning facilities available to young people today, this transition is so typically a blind, trial and error process.

From the standpoint of nomothetic science, the study leaves much to be desired methodologically. Although the authors make appropriate cautionary statements at many points, they cheerfully go on to assertions and conclusions of which they seem quite convinced. For instance, they point out that it is very difficult to distinguish between a coincidence and a correspondence in selected anecdotal materials, yet they find many correspondences and recognize no coincidences. They say, conservatively, their method cannot truly *predict,* yet they conclude that adult adjustment was quite well predicted by the adolescent data. The prediction of adult mental health is summarized in the statistical term of 0.54 (rank correlation). The discussion of the hazards of estimating mental health and the problem of selecting criteria of good mental health, is one of the solid contributions of the book. Personality is defined as the precipitate of the working out or resolution of fantasy; this definition perhaps both narrows the scope of personality and makes undue claims for fantasy. The authors conclude (p. 193) that fantasy "has predictive value for later personality, but that value is uncertain."

The danger in treating such material as scientific data is well illustrated by the following: The authors successfully match thematic protocols taken in adolescence with protocols taken from the same persons 13 yr. later. Early in the volume they note (p. 45) that "actually the formal factors (of expression) were more reliable indicators on which to base matching. Apparently manners of style, language, special comments, fidelity to the picture, emotional tone, and the like are somewhat more prominent characteristics of an individual than are his fantasy themes." Later, this fact—that formal characteristics of language and expression are the principal cues in matching protocols—is forgotten: "There is a marked persistence of themes and fantasy over the 13-yr. interval between this study and its predecessor, enough to make it possible to identify the narrator and match stories told 13 yr. apart when they have been mingled indiscriminately with the stories of other narrators" (p. 209). The wish is father to the thought.

Pennsylvania State University DALE B. HARRIS

Glimpses from the World of the School Child: Self and Other Explored in Dyadic Communication. By AKE BJERSTEDT. New York, Beacon House, 1961, Pp. 131. $3.50.

What kinds of questions would you ask of a stranger if you wished to learn as many important facts as possible about him? What would you tell a stranger about yourself if you wished him to perceive you accurately? These two questions are the major tools with which Bjerstedt obtained data concerning children's self-perceptions and their perceptions of other people. Additional information was obtained by having pairs of children interview one another after each had written his answers to these two critical questions.

The material obtained through the use of these simple techniques is subjected to an exacting content analysis in which "information units" are assigned to niches in content matrices. The rules which govern the coding of information, and the category system itself, are too complicated to be described here, but it should be said that the emphasis is upon the manifest rather than the inferred meaning of subjects' behaviors. Although Bjerstedt acknowledges a debt to Murray for some of the concepts with which he approaches the task of analyzing his protocols, neither the content categories Bjerstedt employed nor the conclusions he reaches reflect great concern about elusive personality processes. This is primarily a book about a technique with which to investigate age and sex differences in the perceptions of children—or of adults. The author generally assumes that the manifest content of subjects' protocols reveals their tendencies to attach differential importance to various people, things, activities, and attributes; but no analysis of underlying personality mechanisms is undertaken.

As might be expected, intercoder reliabilities tend to be high. But the validity of Bjerstedt's interpretations of his data remains in doubt, for he reports no systematic attempt to appraise the correctness of his conclusions. Probably this omission is not very serious because Bjerstedt tends to impute very little surplus meaning to his content categories, and he rarely editorializes about the possible implications of his data. Indeed, some readers will probably criticize him for failure to enumerate the theoretical questions for which his findings might be assumed to provide answers. This book does not attempt to rewrite anybody's theories or to prove that vast stores of new knowledge are soon to revolutionize our thinking. Instead it is content to report that fifth graders, less often than children in lower or higher grades, report facts about themselves or their families when they are asked to write a short essay about themselves. Seventh-grade girls make the most references to their parents. As children progress from the third to the seventh grade there are fewer references to ownership relations and more references to personal relations. The reader is allowed the freedom to incorporate these and many similar findings into his own grand design; apparently Bjetstedt does not care how his axe shall be ground.

The author has described a relatively simple but time-consuming technique by which it is possible to obtain and to categorize data about certain interesting and possibly critical behaviors of children. Presumably these data represent facets of interpersonal and intrapersonal perception. In view of the difficulties which have generally beset scholars in these areas of research, Bjerstedt's efforts should be regarded as a welcome contribution.

University of Illinois I. D. STEINER

A Handbook of Emotional Illness and Treatment. By RICHARD C. ROBERTIELLO. Larchmont, New York, Argonaut Books, 1961. Pp. 159. $3.95.

"The purpose of this book," notes the author is his Introduction, "is to outline for the interested layman and the professional the various kinds of emotional problems and psychiatric illnesses and their manifestations, to indicate their causes and show what can be done about treating them; and, through this, to enable the reader to obtain a better understanding of himself and of the people around him." Section I of the book describes and illustrates the various categories of mental illness, their causes, treatment, and prognosis. Section II consists of a glossary of significant psychiatric terms. Section III tells the lay reader how to determine whether he needs psychotherapy. Section IV outlines various forms of psychiatric treatment. And the fifth and final section tells the reader how to go about seeking psychiatric help.

It is questionable whether professional readers will derive anything of value from this simply and clearly written book; but most of its other material is relevant for, and will probably prove interesting to, the lay reader who wants to learn more about emotional disturbance and its treatment. Unfortunately for this lay reader, Dr. Robertiello's psychoanalytic biases are constantly showing with a vengeance. He seems to acknowledge only two basic forms of psychotherapy: psychoanalysis or psychoanalytic treatment, on the one hand, and what he vaguely calls supportive therapy, on the other hand. Under the former category, he seems to lump all insight-producing treatment, from the orthodox Freudian to the Reichian, Existentialist, and Zen-Buddhist approaches. Under the supportive category, he seems to include widely different schools of therapy, such as rational-persuasive, learning theory, directive, and non-directive schools. The Meyerian psychobiological group he somehow manages to put in both categories. And the fact that most of the analytic groups emphasize relationship therapy, while most of the schools he labels as being "supportive" largely emphasize their patients' obtaining insight into their basic philosophies of living is almost completely ignored by Dr. Robertiello. In a word, his classification and understanding of the various forms of modern psychotherapeutic treatment constitute one big mess.

To make matters considerably worse, although the author is entitled to his psychoanalytic biases, and might well be expected to state that psychoanalytic therapy is the *best* or the *main* method of treating certain emotional disorders, he is hardly entitled to say, as he continually does in this book, that "permanent resolution of the conflict which produces the symptom [of conversion hysteria] can *only* be brought about through psychoanalysis or psychoanalytic psychotherapy" (pp. 53-54); that "treatment of dissociative reaction is *strictly* by psychoanalysis or psychoanalytic psychotherapy" (p. 52); that "the *only* treatment of obsessive-compulsive neurosis is a full-scale psychoanalysis" (p. 56); etc., etc. (italics added by the reviewer). Because of such rash, highly unscientific, and almost certainly false statements as these, as well as because of Dr. Robertiello's serious confusion about what non-psychoanalytic psychotherapy is, I would say that it is the duty of every psychologist to see that this book (which otherwise contains some useful information for lay readers) is *not* placed in the hands of anyone who is interested in obtaining treatment for himself or his associates.

New York City ALBERT ELLIS

The Hidden Springs: An Inquiry into Extra-Sensory Perception. By RENEE HAYNES. New York, The Devin-Adair Company, 1961. Pp. 264. $5.00.

The Hidden Springs contains chapters on such topics as "Organised Divination"; "Psi and Miracle"; "Philosophical Assessments from the World Soul to Prophecy"; together with an appendix on "The Evil Eye." Very little is said about experimental work on ESP. Coover's experiments at Stanford are mentioned but the reader is not acquainted with the fact that they provided no support for ESP. We are told that attempts were made to deny the efficiency with which Dr. J. B. Rhine's experiments were conducted and that "New series of experiments, carried out with all possible precautions, have, however, yielded similar results. . . ." No details of these experiments are given. A great deal of historical and anecdotal material is provided covering such topics as egg-sexing, poltergeists and stigmatization, but those who have read Coover's "Metapsychics and the Incredulity of Psychologists"[1] will find Mrs. Haynes's account of the famous Fox sisters lamentably incomplete. No mention is made of their long history of exposures or of the fact that they made a full public confession in which they stated that all the phenomena with which they were associated were faked.

University of Manchester, England C. E. M. HANSEL

Beyond the Reach of Sense: An Inquiry into Extra-Sensory Perception. By ROSALIND HEYWOOD. New York, E. P. Dutton and Co., 1961. Pp. 224. $3.95.

Any new book on the topic of extra-sensory perception is likely to be of interest to psychologists only in so far as it can throw further light on the experimental evidence for this hypothetical process. Some twenty-five years ago members of the American Psychological Association laid down standards to be observed in experiments in this field: standards that have never yet been observed. In 1939, J. L. Kennedy reviewed the experimental work in the *Psychological Bulletin* and passed only four experiments as being inexplicable in terms of the known forms of experimental error. In 1940, five members of the Parapsychology Laboratory reviewed all the experiments up to that time and came to much the same conclusion as Dr. Kennedy. It is, therefore, tedious to see early experiments, condemned by both critics and supporters of *ESP* alike, resuscitated and cited as providing conclusive evidence.

In this book, the emphasis is on experiments carried out in Britain by members of the Society for Psychical Research. The account of the experimental work is superficial and extremely uncritical. The experiments selected for discussion are, in the main, those that are no longer taken seriously by the majority of parapsychologists. It is, in fact, doubtful whether many of them have heard of Ossowiecki, a Polish engineer who in 1933 guessed the contents of a sealed envelope under conditions reminiscent of a simple conjuring trick. But that experiment is the only one that is reported in any detail.

---

[1] John E. Coover, in *The Case For and Against Psychical Belief,* Ed. by Carl Murchison, Clark University Press, 1927.

The history is given of that keen bunch of amateurs who formed the British Society for Psychical Research as seen through the rose-tinted spectacles of one of its present council members. We learn nothing, however, about the many blunders that formed part of that history.

University of Manchester, England C. E. M. HANSEL

Effective Study. By FRANCES P. ROBINSON. *Rev. ed.* New York, Harper and Brothers, 1961. Pp. ix, 278, $4.50.

The revision of this well-known book on how to study at the college level comes 15 years after the first edition (see this JOURNAL, 60, 1947, 665-666) which was, in turn, a revision of Robinson's first book on studying, *Diagnostic and Remedial Techniques for Effective Study*.

The popularity of *Effective Study* is due principally to Robinson's *SQ3R* Method which is a systematic procedure for studying a textbook. Using this method, a student first *surveys* (skims) a chapter; second, he turns typographical headings into *questions;* third, he *reads* to answer the self-formulated questions; fourth, he *recites* from memory the answers found through reading; and, finally, he *reviews* to end the study period with an over-all mental picture of the chapter.

The major changes in this edition are three: first, and most important, is the addition of one new chapter in which Robinson modifies the *SQ3R* method so that it may now be used for studying non-textbook material, collateral material, English literature, graphs, tables, diagrams, maps, and drawings. This new chapter is entitled, "Other Applications of the *SQ3R* Method of Study."

For the second important change, Robinson clarifies the organization of his book by constructing two additional chapters made up of reworked materials which were previously appended to chapters and embedded in appendices. These new chapter titles are: "Motivation to Study" and "Foreign-Language Study."

The third change deals with one of the book's aims: "This is not a program of reading but of practice." Robinson hopes to achieve this aim by placing at the ends of chapters and in the appendices many new quizzes, tests, questionaries, and exercises; thus, providing the student with an opportunity to begin practicing the principles just read.

In addition to the three major changes just mentioned, Robinson has added three new sections of a functional nature, making the book easier to use. These new sections are: a Bibliography, an Index, and an Appendix which consolidates the quizzes for the various chapters.

There is still another category of change—change making for greater clarity. First, the liberal use of transitions has made quite plain the sequence and relationship of ideas; second, additional and expanded headings and subheadings not only improve readability, but provide the reader with the opportunity to practice one of the main techniques of the *SQ3R*—converting headings into questions; third, the format of the page has been changed from two columns to one; fourth, the better quality of paper, and the addition of sketches and pictures create a favorable attitude; and finally, over 60 percentage of the former footnotes have been replaced with current research studies.

Though some of the chapters (Mathematics, Foreign-Language Study) are relatively below the high quality and practicality of the first six chapters, it is, in the final analysis, Robinson's *SQ3R* Method which distinguishes this book in the area of study skills.

Cornell University WALTER PAUK

National Character and National Stereotypes. By H. C. J. DUIJKER and N. H. FRIJDA. New York, The Humanities Press, 1961. Pp. xi, 238. $4.50.

This is a trend report on recent research, prepared by two psychologists from the University of Amsterdam for the International Union of Scientific Psychology. It is published as the first volume in a new series entitled "CONFLUENCE" edited by the International Committee for Social Sciences Documentation. There is a bibliography of 988 items, taking up 50 pages, with wide international coverage.

The book contains an extensive analysis of the meanings that have been attached to the concept of national character and related concepts in the field of personality and culture. The authors appear to favor the "modal personality" approach to national character. The discussion of methodology is well-considered and justifiably critical of much of the research in this field, particularly because of failure to undertake controlled comparative analyses.

No attempt is made to draw conclusions from the available research as to the psychological characteristics of different nations. However, national stereotypes—what might be termed the layman's ideas about national characteristics as distinguished from the social scientist's ideas—are considered to be generally false. This is not surprising in view of the origin and connotations of the concept of "stereotypes." But it would be more satisfying if the conclusion were based on empirical evidence of the existence or non-existence or particular traits rather than upon the somewhat *a priori* arguments and—let us confess it—moralistic considerations that seem to influence judgment in this murky area.

Few will deny the authors' remarks that the study of national character is not in a very happy state. One may not entirely agree with them, however, as to what is wrong and what needs to be done about it. In the reviewer's opinion, for example, a much harder look needs to be given to the concept of modal personality, with its implication that the different outstanding traits in a given country necessarily coalesce in the same individuals; and much more attention needs to be given to the relation of psychological traits to the political power structure and the class system (for example, is military aggressiveness a trait of national character in a feudal type of society that frequently wars on its neighbors but in which the majority of the people are obedient to their leaders rather than personally aggressive?). The authors pay considerable attention to cultural determination, following the literature on culture and personality, but dismiss abruptly the question of possible somatic determination and pay no heed to possible economic determination. Their consideration of culture does not take account of the existence of "dual cultures" (traditional and modern) in many contemporary nations and the fact of rapid change and transition. Considerations based on the study of culture and personality in small, remote and static folk societies may appear somewhat unrealistic in the present world.

There are some rather bad errors in the references, resembling errors one gets in studies prepared through bureaucratic machinery.

United Nations,
Geneva, Switzerland DONALD V. MCGRANAHAN

Child Development: Volume 1. A Systematic and Empirical Theory. By SIDNEY W. BIJOU and DONALD M. BAER. New York, Appleton-Century-Crofts, 1961. Pp. xii, 90. $1.25.

Many a reader will be surprised at the contents of this slim volume entitled *Child Development,* most especially perhaps the one to whom it is addressed: "the college student who is interested in child behavior and development and has little or no background in psychology" (pp. v-vi). This naive reader may well be astonished at how his interest in children has led him directly to a study of such phenomena as operant stimulus generalization, the temporal gradient of reinforcement, discriminative stimuli, schedules of reinforcement, and the like. Fewer eyebrows will be raised by the authors' professional colleagues, familiar with their program of research within the Skinnerian tradition.

The modestly priced paperback is the first in a projected series of four. The three forthcoming will be concerned with infancy, childhood, and adolescence. The present volume sets the theoretical context for these later ones: "The developing child is adequately conceptualized as a source of responses which fall into two functional classes: responses . . . and operants . . ." (p. 84). On its own, it should find wide use as a supplementary text in courses in developmental and educational psychology. It presents a clear and straightforward—if somewhat heavyhanded—account of development as learning, occurring through selective reinforcement of respondents and operants. Through it, students will see connections, perhaps hitherto unsuspected, between the field of child psychology and the contents of their introductory psychology course.

To most psychologists, the special appeal of the Skinnerian approach lies in its close devotion to observable responses and to the experimental method of studying functional relations. The authors' enthusiasm has the same basis; they argue for a natural science point of view, emphasizing observables. It seems unfortunate that this volume's readers are given no opportunity here to experience the natural science feature of the approach, no exposure to the data. We may hope that in the succeeding volumes the authors will press their advantage, utilize the special appeal of their view, by displaying the experimental evidence for the theory they present and its applicability of human development. In deference to the level of preparation of their readers, here they have been content to buttress their account of development only by alluding to anecdotal evidence and by citing the standard Skinnerian texts, none of which is distinguished by special attention to the human child.

Center for Advanced Study
in the Behavioral Sciences

ALBERTA ENGVALL SIEGEL

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages.
Listing here does not, however, preclude their later review.)

ACKOFF, RUSSELL L. (with the collaboration of SHIV K. GUPTA and J. SAYER MINAS), *Scientific Method: Optimizing Applied Research Decisions.* New York, John Wiley & Sons, 1962. Pp. xii, 464. $10.25.

ALEXANDER, ARTHUR, *The Hidden You: Psychology in Your Life.* Englewood Cliffs, N.J., Prentice-Hall, 1962. Pp. 63. $2.95.

ALTSCHUL, A., *Aids to Psychology for Nurses.* London, Baillere, Tindall & Cox, 1962 Pp. xii, 308. $3.00 (paperback).

Armstrong, D. M., *Bodily Sensations.* New York, Humanities Press, 1962. Pp. vii, 132. $2.50.

Bachrach, Arthur J., *Psychological Research: An Introduction.* New York, Random House, 1962. Pp. x, 113. $1.45 (paperback).

Bauer, Raymond A., (ed.), *Some Views on Soviet Psychology.* Washington, D.C., American Psychological Assoc., 1962. Pp. ix, 285.

Berg, Charles, *Madkind: The Origin and Development of the Mind.* New York, Hillary House, 1962. Pp. 277. $7.00.

Blair, G. M., Jones, R. S., and Simpson, R. H., *Educational Psychology,* 2nd ed. New York, Macmillan Co., 1954 and 1962. Pp. xxiv, 678. $7.00.

Brehm, J. W., and Cohen, A. R., *Explorations in Cognitive Dissonance.* New York, John Wiley, 1962. Pp. xiv, 334. $7.95.

Bricklin, B., Piotrowski, Z. A., and Wagner, E. E., *The Hand Test.* Springfield, Ill., Charles C Thomas, 1962. Pp. x, 100. $5.00.

Brown, R., Galanter, E., Mandler, G., and Hess, E. H., *New Directions in Psychology.* New York, Holt, Rinehart & Winston, 1962. Pp. ix, 353. $2.50. (paperbound)

Chappell, V. C., (ed.), *The Philosophy of Mind.* Englewood Cliffs, N.J., Prentice-Hall, 1962. Pp. ix, 178. $2.25. (paperback)

Coon, Carleton S., *The Origin of Races.* New York, Knopf, 1962. Pp. xli, 724, xxi. $10.00.

Crombie, I. M., *An Examination of Plato's Doctrines: 1. Plato on Man and Society.* New York, Humanities Press, 1962. Pp. xii, 395. $8.50.

Dingwall, Eric J., *Some Human Oddities.* New Hyde Park, N.Y., Univ. Books, Inc., 1962. Pp. 198. $6.00.

Dreger, Ralph Mason, *Fundamentals of Personality.* Philadelphia, J. B. Lippincott Co., 1962. Pp. xi, 388. $6.95.

Dyal, James A., *Readings in Psychology: Understanding Human Behavior.* New York, McGraw-Hill, 1962. Pp. 444. $3.95. (paperback)

Eiseley, Loren, *The Mind as Nature.* New York, Harper & Row, 1962. Pp. 60. $2.75.

Ennis, Merlin (collected and translated), *Umbundu: Folk Tales from Angola.* Boston, Beacon Press, 1962. Pp. xxix, 316. $7.50.

Eriksen, C. W., (ed.), *Behavior and Awareness: A Symposium of Research and Interpretation.* Durham, N.C., Duke Univ. Press, 1962. Pp. xi, 158. $3.50.

Estabrooks, G. H., (ed.), *Hypnosis: Current Problems.* New York, Harper & Row, 1962. Pp. viii, 285. $5.50.

Farnsworth, P. R., (ed.), *Annual Review of Psychology.* Vol. 13 Palo Alto, Calif., Stanford Univ. Press, 1962. Pp. ix, 602. $7.00.

Feigl, Herbert, and Maxwell, Grover, (eds.), *Scientific Explanation, Space, and Time. Minnesota Studies in the Philosophy of Science, Vol. 111.* Minneapolis, Minn., Univ. of Minnesota Press, 1962. Pp. xv, 628. $8.50.

Fuller, John L., *Motivation: A Biological Perspective.* New York, Random House, 1962. Pp. vi, 116. $1.45. (paperback)

Galton, Francis, *Hereditary Genius.* New York, Meridian Books, Inc., 1962. (First published 1869) Pp. 446. $1.45. (paperback)

Garre, Walter J., *Basic Anxiety: A New Psychobiological Concept.* New York, Philosophical Library, 1962. Pp. 123. $5.00.

Garrett, J. F., and Levine, E. S., (eds.), *Psychological Practices with the Physically Disabled.* New York, Columbia Univ. Press, 1962. Pp. xi, 463. $8.75.

Geldard, Frank A., *Fundamentals of Psychology.* New York, John Wiley & Sons, 1962. Pp. 437. $7.50.

Gendlin, Eugene T., *Experiencing and the Creation of Meaning: A Philosophical and Psychological Approach to the Subjective.* New York, Free Press of Glencoe, 1962. Pp. xv, 302. $6.00.

Gordon, Ira J., *Human Development: From Birth through Adolescence.* New York, Harper & Bros., 1962. Pp. xiii, 400. $5.75.

Gorow, Frank F., *Statistical Measures: A Programmed Text.* San Francisco, Calif., Chandler, 1962. Pp. viii, 178. $2.25. (paperback)

GREEN, EDWARD J., *The Learning Process and Programmed Instruction.* New York, Holt, Rinehart & Winston, 1962. Pp. x, 228. $4.00.

HAMMOND, KENNETH R., and HOUSEHOLDER, JAMES, *Introduction to the Statistical Method.* New York, Knopf, 1962. Pp. xvi, 412, vii. $7.00.

HART, BERNARD, *The Psychology of Insanity.* New York, Cambridge Univ. Press, 1962 (reprint of fifth edition, 1957). Pp. xi, 127. $1.25. (paperback)

HAWKES, GLENN R., and PEASE, DAMARIS, *Behavior and Development from 5 to 12.* New York, Harper & Bros., 1962. Pp. xiv, 375. $5.75.

JONES, MARSHALL R., (ed.), *Nebraska Symposium on Motivation. Vol. 10.* Lincoln, Neb., Univ. of Nebraska Press, 1962. Pp. xiii, 330. $3.25 paper, $4.25 cloth.

KAGAN, JEROME, and MOSS, HOWARD A., *Birth to Maturity: A Study in Psychological Development.* New York, John Wiley & Sons, 1962. Pp. xiii, 381. $8.50.

KAHN, SAMUEL, *Thanks for a Better Memory.* New York, Vantage Press, 1962. Pp. 176. $3.00.

KAPLAN, BERNARD, and WAPNER, SEYMOUR, (eds.), *Perspectives in Psychological Theory: Essays in Honor of Heinz Werner.* New York, International Universities Press, 1960. Pp. 384. $7.50.

KARN, HARRY W., and GILMER, B. VON HALLER, *Readings in Industrial and Business Psychology,* 2nd ed. New York, McGraw-Hill, 1962. Pp. xi, 515. $4.95 paper, $6.95 cloth.

KLEIST, K., *Sensory Aphasia and Amusia.* New York, Pergamon Press, 1962. Pp. xv, 95. $5.00.

LAZARUS, RICHARD S., *Personality and Adjustment.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. x, 118. $3.95 cloth, $1.50 paper.

MADIGAN, MARIAN EAST, *Psychology: Principles and Applications,* 3rd ed. St. Louis, Mo., C. V. Mosby Co., 1962. Pp. 379. $5.50.

MCNEMAR, QUINN, *Psychological Statistics,* 3rd ed. New York, John Wiley & Sons, 1962. Pp. vii, 451. $7.75.

MILLER, GEORGE A., *Psychology: A Science of Mental Life.* New York, Harper & Row, 1962. Pp. xi, 388. $6.95.

MILLS, ENID, *Living with Mental Illness: A Study in East London.* New York, Humanities Press, 1962. Pp. xii, 184. $5.50.

MUNN, NORMAN L., *Introduction to Psychology,* 4th ed. Boston, Houghton Mifflin, 1962, Pp. xiv, 588. $6.95.

NASH, HARVEY, *Alcohol and Caffeine. A Study of their Psychological Effects.* Springfield, Ill., Charles C Thomas, 1962. Pp. xi, 169. $7.25.

NIELSEN, GERHARD S., (ed.), *Proceedings of XIV International Congress of Applied Psychology. Vol. 1. Psychology and International Affairs: Can We Contribute?; Vol. 2. Personality Research; Vol. 3. Child and Education; Vol. 4. Clinical Psychology; Vol. 5. Industrial and Business Psychology.* Copenhagen, Munksgaard, 1962. Pp. 132, 229, 197, 211, 229.

NIXON, ROBERT E., *The Art of Growing: A Guide to Psychological Maturity.* New York, Random House, 1962. Pp. xxxi, 159. $3.95.

NOLL, VICTOR H., and NOLL, RACHEL P., *Readings in Educational Psychology.* New York, Macmillan Co., 1962. Pp. xix, 568. $3.95. (paperback)

NUTTIN, JOSEPH, *Psychology in Belgium.* Louvain, Publications Universitaires, 1961. Pp. 80. FB 70. ($1.60).

O'NEIL, W. M., *An Introduction to Method in Psychology,* 2nd ed. New York, Melbourne Univ. Press, Cambridge Univ. Press, 1962. Pp. x, 178. $4.50.

PARNES, SIDNEY J., and HARDING, HAROLD F., (eds.), *A Source Book for Creative Thinking.* New York, Charles Scribner's Sons, 1962. Pp. xiv, 393. $4.75 paper, $6.50 cloth.

PAUK, WALTER, *How to Study in College.* New York, Houghton Mifflin Co., 1962. Pp. viii, 132. $1.95. (paperback)

PAVLOV, I. P., *Essays in Psychology and Psychiatry with a Special Section on Sleep and Hypnosis.* New York, Citadel, 1962 (copyright 1957). Pp. 189. $1.95. (paperback)

PLUTCHIK, ROBERT, *The Emotions: Facts, Theories, and a New Model.* New York, Random House, 1962. Pp. viii, 204. $1.95. (paperback)

POSTMAN, LEO, (ed.), *Psychology in the Making: Histories of Selected Research Problems.* New York, Alfred A. Knopf, 1962. Pp. 785, xxxiv.

RAY, WILLIAM S., *Statistics in Psychological Research.* New York, Macmillan Co., 1962. Pp. xi, 303. $6.00.

REICHARD, SUZANNE, LIVSON, FLORINE, and PETERSEN, P. G., *Aging and Personality: A Study of 87 Older Men.* New York, John Wiley & Sons, 1962. Pp. xv, 237. $7.95.

RIPPLE, RICHARD E., and KLAUSMEIER, HERBERT J., *Adjunct Program to Accompany Learning and Human Abilities: Educational Psychology.* New York, Harper & Row, 1962. Pp. xxi, 336. $3.50. (paperback)

SARTAIN, A. Q., NORTH, A. J., STRANGE, J. R., and CHAPMAN, H. M., *Psychology: Understanding Human Behavior,* 2nd ed. New York, McGraw-Hill, 1962 (1958), Pp. xi, 432. $6.75.

SCHAEFER, KARL E., (ed.), *Environmental Effects on Consciousness.* New York, Macmillan Co., 1962. Pp. x, 146. $5.50.

SCHWARZ, BERTHOLD E., and RUGGIERI, BARTHOLOMEW A., *Parent-Child Tensions.* Philadelphia, J. B. Lippincott, 1958. Pp. xv, 238. $4.95.

SIMMONS, JAMES R., *The Quest for Ethics.* New York, Philosophical Library, 1962. Pp. 54. $2.75.

SMITH, HENRY P., *Psychology in Teaching,* 2nd ed. Englewood Cliffs, N.J., Prentice-Hall, 1962. Pp. xiv, 498. $6.95.

SUNDBERG, NORMAN D., and TYLER, LEONA E., *Clinical Psychology: An Introduction to Research and Practice.* New York, Appleton-Century-Crofts, 1962. Pp. xviii, 564. $7.00.

TALLENT, NORMAN, *Clinical Psychological Consultation: A Rationale and Guide to Team Practice.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. xviii, 298. $8.65 trade, $6.50 text.

TASSI, DAN, *The Mind and Time and Space.* Philadelphia, Dorrance & Co., 1962. Pp. 114, $3.00.

THORPE, LOUIS P. (in collaboration with VIRGINIA JOHNSON), *Child Psychology and Development,* 3rd ed. New York, Ronald Press, copyright 1946, 1955, 1962. Pp. vii, 628. $7.00.

TIBBITTS, CLARK, and DONAHUE, WILMA, *Social and Psychological Aspects of Aging.* New York, Columbia Univ. Press, 1962. Pp. xviii, 952. $20.00.

TOWNSEND, E. A., and BRUKE, PAUL J., *Learning for Teachers.* New York, Macmillan Co., 1962. Pp. vi, 313. $5.50.

TYLER, LEONA E., *Tests and Measurements.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. xi, 116. $3.75 text, $5.00 trade (paperback).

VILLINGER, W., and STUTTE, H., *Jahrbuch für Jugen-Psychiatrie und ihre Grenzgebiete.* Vol. III. Bern and Stuttgart, Hans Huber, 1962. Pp. 299. Fr./DM 32.-

VON GREYERZ, W., *Psychology of Survival: Human Reactions to the Catastrophes of War,* New York, American Elsevier Publishing Co., Inc., 1962. Pp. 99. $3.75.

WARFEL, HARRY R., *Language: A Science of Human Behavior.* Cleveland, Ohio, Howard Allen, 1962. Pp. 188. $3.95.

WOLFGANG, M. E., SAVITZ, L., and JOHNSTON, N., (eds.), *The Sociology of Crime and Delinquency.* New York, John Wiley & Sons, 1962. Pp. xiv, 423. $4.45 paper, $6.75 cloth.

WOOLF, HARRY, (ed.), *Quantification.* New York, Bobbs-Merrill, 1961. Pp. 224. $6.50.

YOUNG, MORRIS N., and GIBSON, WALTER B., *How to Develop an Exceptional Memory.* Philadelphia, Chilton Books, 1962. Pp. vii, 265. $4.95.

THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. Stanley Hall

Vol. LXXVI JUNE, 1963 No. 2

THE SHAPE-DISCRIMINATION OF STATIONARY SHAPES BY OCTOPUSES

By N. S. Sutherland, Oxford University, England

The present experiment was primarily undertaken to answer two questions. (1) Can octopuses discriminate stationary shapes? If they can, do they (2) discriminate between stationary and moving shapes in terms of the same properties? Two groups of these animals were trained to discriminate between the same shapes, one group with the shapes stationary, the other with the shapes in motion. Transfer-tests with 30 further shapes then were given. If both groups react to the transfer shapes in the same way, we can reasonably infer that they performed the original discrimination by analyzing the shapes in terms of the same properties. *

The training and transfer-shapes used were so selected that the results would provide a test of some predictions made from a theory of shape-discrimination in octopuses. It has been suggested that octopuses discriminate certain shapes by analyzing their vertical and horizontal 'projections'.[1] The vertical projection of a shape is obtained by counting its vertical extent at each point on the horizontal axis, and the horizontal projection by counting its horizontal extent at each point on the vertical axis. Since there is some evidence that horizontal differences govern octopuses' behavior more readily than do vertical differences, it follows that shapes differing only in their horizontal projections should be more readily discriminable than shapes differing only in their vertical projections.[2] The prediction was not

---

* Received for publication December 21, 1961. This research was supported in part by the U. S. Office of Naval Research under terms of Contract N62558-2453, and in part by the Nuffield Foundation. I am indebted to Peter Dohrn, Director of Stazione Zoologica, Naples, for the facilities afforded me there; to A. E. Carr, J. A. Mackintosh, and J. L. Sutherland for help in the conduct of the experiments; and to H.-L. Teuber and J. Z. Young for many valuable suggestions made while the paper was in preparation.

[1] N. S. Sutherland, Visual discrimination of orientation and shape by *Octopus*, *Nature,* 179, 1957, 11-13; Theories of shape discrimination in *Octopus*, ibid., 186, 1960, 840-844.

[2] Sutherland, Visual discrimination of orientation by *Octopus:* Mirror images, *Brit. J. Psychol.,* 51, 1960, 9-18; Discrimination of horizontal and vertical extents by *Octopus*, *J. comp. physiol. Psychol.,* 54, 1961, 43-48.

however, confirmed by the results of an experiment with the two pairs of shapes shown at the top of Fig. 1; the two pairs of shapes were equally discriminable.[3] The results of transfer-tests suggested that the octopus might employ a second shape-analyzing mechanism based on the ratio $C/A^{\frac{1}{2}}$ for each shape. Shapes high on this ratio will be called 'open' shapes, those low on this ratio 'closed' shapes.

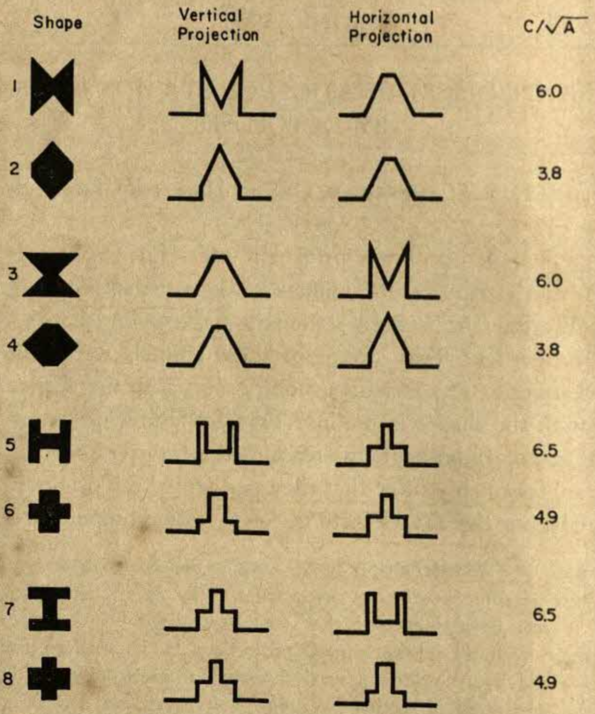


FIG. 1. VERTICAL AND HORIZONTAL PROJECTIONS OF OPEN AND CLOSED SHAPES In each pair the open shape is first, the closed second. The training-shapes used in this experiment are Pair 5-6 (vertical shapes) and Pair 7-8 (horizontal shapes).

The present experiment attempts to confirm and extend the findings of the previous one on open and closed forms. The *Ss* were trained on two new pairs of shapes differing either in horizontal projections alone (Fig. 1, Shapes 5 and 6) or in vertical projections alone (Shapes 7 and 8). These shapes have a higher $C/A^{\frac{1}{2}}$-ratio than those previously used. In the earlier experiment, with training-shapes of ratios of 3.8 and 6.0, there was no transfer to shapes with ratios of 4.5-5.0. In the present experiment, such transfer-shapes should be treated as equivalent to the original closed shape (ratio of 4.9), and a test of the hypothesis that octopuses may analyze shapes in terms of the ratio $C/A^{\frac{1}{2}}$ thus is provided.

---

[3] Sutherland, Visual discrimination of shape by *Octopus:* Open and closed forms, *J. comp. physiol. Psychol.*, 53, 1960, 104-112.

Method: (1) Subjects. The *S*s were 16 octopuses (*Octopus vulgaris* Lamarck), weighting from 250-500 gm. Each was kept in an individual tank with a 'home' of bricks at one end. Eight of the animals (Group M) were trained with moving shapes, and eight (Group S) with stationary shapes.

(2) Experimental procedure. The training procedure used for Group M has been described fully elsewhere.[4] Details will be given here only of the procedure used for Group S.

(3) Method of presenting shapes. All shapes were cut from ¼-in. white Perspex and were attached to transparent Perspex rods.[5] Before each trial with *S*s of Group S, an opaque gray formica screen was placed in position between *S* and the wall of the tank at the far end from its home. The top of the screen fitted flush with the half-opened lid of the tank (see Fig. 2). The shape was then introduced behind the screen without *S*'s seeing it. The rod carrying the shape was supported by a transparent Perspex bar whose ends rested on the sides of the tank. When in position, the center of the shape was about 15 cm. from the bottom of the tank and 8 cm. in front of the far wall. Each trial was begun by removing the opaque screen, thus enabling *S* to see the stationary shape. No screen was used with Group M, and the shapes were kept in motion while in the water.

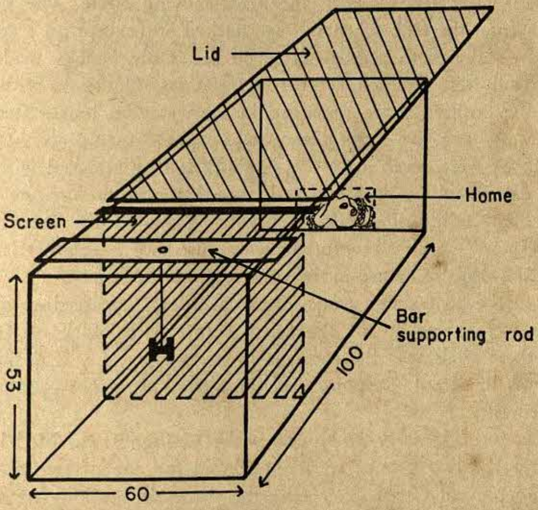


FIG. 2. TRAINING SITUATION USED WITH STATIONARY SHAPES

(4) Pretraining. All *S*s first were trained to attack pieces of sardine, then to attack the positive shape. If an animal of Group S failed to attack the stationary

---

[4] Sutherland, *op. cit.*, *J. comp. physiol. Psychol.*, 54, 1961, 43-48.

[5] Except for two of the transfer-shapes which were cut from black Perspex.

shape within 30 sec. of its being exposed, a moving black circle was inserted near the shape. If *S* then attacked, the circle was removed before it could be grasped; invariably, then, *S* would grasp the stationary positive shape. If *S* failed to attack even when the moving circle was shown, it was lured out to the stationary shape with a piece of sardine. Any attack on the positive shape was rewarded with a piece of sardine weighing about 2.5 gm. Group S received 150 pretraining trials with the positive shape; 10 trials were given per day. Group M was trained to attack the moving positive shape; only 20 trials were given.

(5) Training. During the experimental training, which began on the day following the last day of pretraining, all *Ss* received 20 trials a day in two blocks of 10 trials each; one block given in the morning, and the other in the early evening. Within each block, the positive and negative shapes were shown five times each in a quasi-random order. The intertrial interval was 10 min.

On positive trials, the animals of Group S were rewarded with a piece of sardine if they attacked the shape within 60 sec. after the screen was raised. If *S* failed to attack within this time on three consecutive positive trials, it was lured out to the shape after 60 sec. with the black moving circle. The positive shape was withdrawn from the tank when it was attacked or (except on a trial following three consecutive failures to attack) after 60 sec. Only attacks made within the first 30-sec. period were counted as correct responses. On trials with the animals of Group M, the positive shape remained in the water no longer than 20 sec.

For Group S, the negative shape always was exposed for 30 sec. after the screen was raised. The *Ss* were given a 10-v. A.C. shock if they attacked within this time. After the first day's training, no *S* attacked more than once on any one trial with the negative shape. Any failure to attack the negative shape within 30 sec. was counted as a correct response. A similar procedure was used on negative trials for Group M, except that the shape remained in the water only for 20 sec.

In addition to the 20 training trials given each day, 2 prefeeding trials with the positive shape were given to all the *Ss*, 1 before each block of 10 trials. The procedure on these trials was the same as in pretraining; *i.e.* if *S* failed to attack the shape, it was lured out. Responses on these trials were not counted in computing training scores.

(6) Comment on differences in procedure for Groups S *and* M. The differences in procedure for the two groups were introduced for one of two reasons: (1) It was necessary to prevent Group S's seeing the shapes in motion—hence the screen; (2) it was desirable to ensure that Group S made about the same total number of attacks as did Group M. Since, however, stationary shapes are less attractive than moving ones, it was necessary (a) to give more pretraining trials to Group S; (b) to count any attack made within 30 sec. as a response for Group S, instead of counting only attacks made within 20 sec.; (c) to allow attacks on the stationary positive shape up to 60 sec., even though such attacks were not counted as correct if they occurred in the second 30-sec. period; (d) to lure the *Ss* of Group S out to the positive shape after 60 sec. if they consistently failed to attack.

(7) Transfer-tests. All of the *Ss* were given transfer-tests after the period of initial training. Transfer-shapes were presented to each *S* in exactly the same way as the training shapes for that *S*, except that attacks were neither rewarded nor punished. Transfer-shapes were removed from the water when attacked or after 30 sec. for Group S (20 sec. for Group M). Twenty transfer-trials were given per

day, and each one was preceded by a retraining trial with the original training shapes. The order in which different transfer-shapes were shown was random except that each was shown once before any was shown again. The order was such that each transfer-shape was shown five times after a preceding positive retraining trial and five times after a preceding negative retraining trial.

(8) Experimental design. Within Groups M and S there were two main subgroups, each of four *S*s. Subgroups V were trained to discriminate shapes differing in their vertical projections (Shapes 5 and 6 in Fig. 1). Subgroups H were trained with shapes differing in their horizontal projections (Shapes 7 and 8 in Fig. 1). Within each subgroup, two *S*s were trained with the open shape positive, two with the closed shape positive. Of each such pair of *S*s from Group M, one was trained with shapes moving up and down, and the other with shapes moving from side to side. All the *S*s received 280 training trials. They then were given 10 transfer-tests with each of the shapes shown in Fig. 4. Group M received additional transfer-tests with four shapes moved at 90° to their original direction of motion; these shapes were the two original training shapes and their 90° rotations.

Results: (1) Pretraining. All the *S*s in Group M readily attacked the moving positive shapes from the outset of pretraining. Since 20 trials were given to each *S*, there was a total of 160 trials, and on only 5 trials did the *S*s fail to attack within 20 sec.

The results for Group S present a very different picture. On the first day, they made 31 attacks within 30 sec. out of 80 trials. There was considerable variability among the *S*s; 2 attacked on 9 out of 10 trials, and 3 failed to attack at all on the first day. Nevertheless, the *S*s which at first did not attack a stationary shape presented alone gradually learned to do so, and on Days 7-9 attacks were made on 83% of the trials. Thereafter, there was little increase in the tendency to attack, and pretraining was terminated on the fifteenth day, by which time the *S*s were attacking on 88% of trials. The latencies of Group S on trials on which attacks were made were much higher than those of Group M. The mean latencies on the last day of pretraining were 9.9 and 4.8 sec., respectively.

(2) Training on stationary shapes. The learning curves for Groups M and S are shown in Fig. 3. It is clear that the *S*s trained with stationary shapes learned almost as quickly as those trained with moving shapes and reached as high an asymptote of performance. It is, however, dangerous to compare the results too closely, since the two groups were trained in different years. The dip in the performance of Group M over Days 5-12 is, in fact, known to have been due to the poor quality of the circulating seawater over this period.[6] What is certain is that the best *S*s in each group reached about the same asymptote of performance—80-90% correct

---

[6] See Sutherland, *op. cit., J. comp. physiol. Psychol.*, 54, 1961, 43-48, for an account of another experiment carried out at this time in which the learning curves showed a similar dip.

—and this level was maintained by these *Ss* throughout the long series of transfer-tests which followed the initial training. Thus it is established that octopuses can learn to discriminate stationary shapes to as high an asymptote of performance as moving shapes. Since the good performers in Group S were attacking on almost exactly 50% of all trials toward the end of training, the methods used for maintaining the level of attack on stationary shapes appear to have been successful.

There were striking qualitative differences in the behavior of the *Ss* from Groups M and S. Those trained with stationary shapes moved out to attack in a slow and deliberate manner; the shape would be seized with one or two arms and drawn slowly toward the animal. This behavior contrasts with the sudden movements made by the *Ss* being trained on moving shapes; they often would hurl themselves onto the shape, or, if they attacked more tentatively, would do so by coming close, stopping, and then suddenly flicking at the shape with fast movements of the arms. After learning had occurred, the animals of Group S often would not leave the home when the negative shape was shown, or they would come out slowly and pause half-way to watch it. Such responses also are made by animals trained with moving shapes.

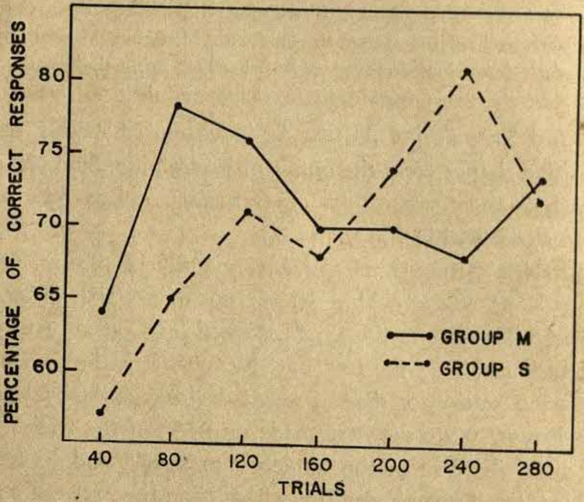


FIG. 3. PERFORMANCE DURING INITIAL TRAINING OF GROUP *M* AND GROUP *S*

(3) Discriminability of the training-shapes. There was no significant difference in performance during training between Groups M-V and M-H; they averaged, respectively, 74% and 70% correct responses over the 280

training trials. By contrast, Group S-H averaged 79% correct over this period and Group S-V only 61%. Although this difference is highly significant ($t = 6.8$, $df = 6$, $p < 0.001$), and although it is in the direction predicted, it must be interpreted with caution. The animals of Group S-V attacked much less frequently during pretraining than did those of Group S-H; three S-V animals attacked less frequently than any S-H animal at all stages of training. If animals do not attack, they have no opportunity to learn, and it seems possible, therefore, that the difference in performance between the subgroups of Group S was due to chance differences in readiness to attack. Further evidence will be adduced below which suggests that this difference should be set down to sampling error.

An analysis of variance performed on training scores showed that no factor was significant other than that already discussed; namely, the interaction between method of training (moving or stationary shapes) and shapes used (H or V). A second analysis of variance performed on Group M revealed that the effects of different directions of movement were not significant, although there was a tendency for the *S*s trained with the open shape moved in the direction of its long arms to perform better (76%) than the *S*s trained with this shape moved at right angles to its long arms (68%); this finding confirms an earlier one which was significant.[7]

(4) Transfer-tests: (A) Method of analysis. The transfer-shapes used are shown in Fig. 4, and Table I summarises the results. The method of analysis can best be grasped by referring to Table I. The numbers in Column 1 refer to the shapes drawn in Fig. 4. Columns 2 and 3 give, respectively, the number of attacks made (out of 80 trials with each shape)[8] by the 8 *S*s trained with the original open-positive and by the 8 trained with the original closed-positive. Columns 4 and 5 are calculated from Columns 2 and 3. Column 4 gives the percentage of trials on which the *S*s treated each transfer-shape as an open one *i.e.* the percentage of trials on which the *S*s trained with open-positive attacked the shape and the *S*s trained with closed-positive failed to attack the shape. Column 5 shows the percentage of all attacks made by the *S*s trained with the open-positive shape. From Columns 4 and 5, it is possible to rank transfer-shapes in terms of their similarity (for octopuses) to the original open and closed shapes. The order in which shapes are numbered in Fig. 4 and in Table I is based on the rank-order obtained from the values in Column 4; where these values are equal, the order was determined from Column 5. Thus, values above 50% in these two columns represent increasing transfer from the open shape, and values below 50% represent increasing transfer from the closed shape. Columns 6-9 give the rank-order of transfer-shapes in terms of decreasing transfer from the original open shape separately for four dif-

---

[7] Sutherland, *op. cit.*, *J. comp. physiol. Psychol.*, 53, 1960, 104-112.

[8] The figures for Shapes *O* and *C* (the original training shapes) represent the *average* number of attacks made per 80 trials during the retraining which was interspersed with transfer-tests.

ferent groups of 8 animals each. The rank-orders were arrived at in the way explained above by considering for each group the number of attacks made on each transfer-shape by the *Ss* trained with the original open-positive shape and by those trained with the original closed-positive.

The reliability of the results is difficult to estimate, since the number of attacks made by the individual *Ss* on different transfer-shapes is determined partly by their general readiness to attack and by how well they were discriminating between the original training shapes as well as by chance-fluctuations in readiness to attack. Rank-order correlations (ρ) were computed between the transfer-ranks obtained from different groups. The correlation-coefficients for Groups V and H and for Groups M and S are, respectively, 0.72 and 0.69. These are based on rankings obtained from groups of 8 *Ss*, and they are higher than any of the correlations between subgroups of 4*Ss* (*e.g.* ρ for Groups M-V and S-V = 0.52), which suggests that the final order based on 16 *Ss* is reasonably reliable. Moreover, 16 of the present transfer-shapes are directly comparable to 16 shapes used in the previous experiment on open and closed forms, and the value of ρ obtained for the correlation between the orderings of these 16 shapes in the two experiments is 0.78.[9] Two further experiments have since been performed involving open and closed forms, and the results are in good agreement with those found here.[10]

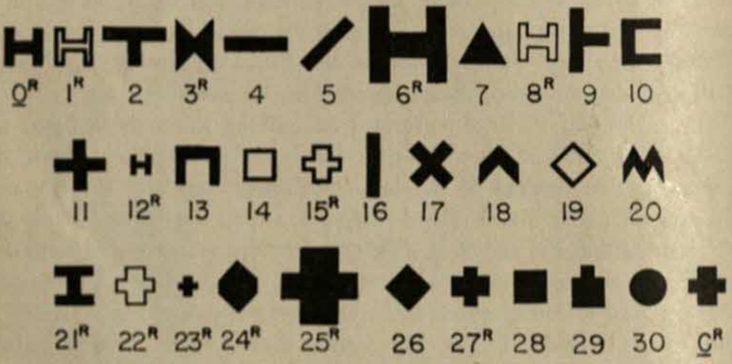


FIG. 4. TRANSFER-SHAPES

Shapes *O* and *C* are the original vertical training-shapes. Shapes are drawn in the orientation in which they were shown to Group *V*. Shapes marked *R* were shown to Group *H* rotated through 90°. The order in which the transfer-shapes are arranged was determined by the relative amount of transfer to each from the original-training shapes.

(b) Results of the transfer-tests. It is worth noting some features of the ordering of transfer-shapes which appear both in this experiment and in

---

[9] Sutherland, *op. cit., J. comp. physiol. Psychol.*, 53, 1960, 104-112.

[10] Sutherland, Visual discrimination of shape by *Octopus:* Squares and crosses, *J. comp. physiol. Psychol.*, 55, 1962, 939-943; N. S. Sutherland, N. J. Mackintosh, and J. A. Mackintosh. Visual discrimination of shape by *Octopus:* Squares and parallelograms (in preparation).

the previous one on open and closed forms, and which would not be expected by common sense. (i) In both experiments there is excellent transfer from the original open shape to an outline of it (No. 1).[11] There is no transfer from the original closed shape to its outline (No. 15). (ii) The upright cross (No. 11) is treated as a more open shape than the oblique cross (No. 17). (iii) A circle (No. 30) is treated as a more closed shape than a square (No. 28). (iv) There is a tendency to treat an equilateral triangle (No. 7) as an open shape, not as a closed shape. (v)

TABLE I

RESULTS OF TRANSFER-TESTS

1 Shape	2 Attacks by *S*s trained with *O*+	3 Attacks by *S*s trained with *O*−	4 % times treated as open trials	5 % attacks by *S*s trained with *O*+/all attacks	6 Group *M* ranks	7 Group S ranks	8 Group V ranks	9 Group H ranks	10 $C/A^{1/2}$	11 Ranks on $C/A^{1/2}$
0	48	9	74	81					6.5	
1	55	15	75	79	1	2	2	1	15.5	1
2	47	17	69	73	5	3.5	12	2	6.1	9.5
3	42	13	68	76	5	5	5	5	6.0	12
4	51	22	68	70	14	1	8.5	4	5.4	17.5
5	40	12	68	77	9	3.5	4	6	5.4	17.5
6	49	21	68	70	3	12.5	2	9	6.5	7
7	41	16	66	72	7.5	8	2	15	4.6	23
8	37	14	64	73	6.5	8	6	7		
9	39	17	64	70	2	20.5	8.5	9	6.1	9.5
10	46	24	64	66	10	8	19	3	5.7	14.5
11	38	20	61	66	12.5	10.5	14	12	6.0	12
12	38	21	61	64	17	6	14	14	6.5	7
13	37	21	60	64	5	20.5	17	12	5.7	14.5
14	37	21	60	64	11	15.5	8.5	17	12.0	3.5
15	31	19	58	62	12.5	18	17	16	13.1	2
16	31	21	56	61	15.5	18	14	19.5	5.4	17.5
17	34	24	56	59	19	14	20.5	12	6.0	12
18	32	22	56	59	21	10.5	17	18	5.4	17.5
19	34	25	56	58	19	15.5	11	21.5	12.0	3.5
20	46	37	56	55	15.5	18	8.5	21.5	7.6	5
21	30	24	63	56	22	12.5	23	9	6.5	7
22	28	30	49	48	19	24.5	22	19.5		
23	25	32	46	44	23	22	20.5	24	4.9	21
24	22	36	41	38	24.5	23	24.5	23	3.8	27
25	24	43	38	36	26.5	24.5	24.5	25	4.9	23
26	15	39	34	28	26.5	26	26	26.5	4.0	24
27	11	40	32	28	29	27	29	26.5	4.9	25
28	17	47	31	27	24.5	29	28	28	4.0	26
29	14	44	31	24	28	28	27	29	4.5	27
30	11	54	23	17	30	30	30	30	3.5	28
C	9	45	28	17					4.9	

---

[11] Here and henceforth, Arabic numerals in brackets refer to transfer-shapes drawn in Fig. 4.

Rectangles (Nos. 4, 5, 16) are treated as open shapes, not as closed shapes. (vi) There is excellent transfer to the rotated closed shape, but there is very little transfer to the rotated open shape.

All the above results were obtained in both experiments. Three further results of the present transfer-tests deserve comment here. (vii) There is some transfer from white to black shapes, though transfer is incomplete (Nos. 8, 22). (viii) Transfer to larger shapes (Nos. 6, 25) is better than transfer to smaller shapes (Nos. 12, 23); this result now has appeared in many different experiments and is fully discussed elsewhere.[12] (ix) Four additional transfer-tests involving shapes moving at right angles to their original direction of motion were given to Group M. The four shapes moved in this way and the percentages of times each was treated like the original open shape were 64% for the original open shape; 40% for the original closed shape; 55% for the 90°-rotation of the original open shape; and 41% for the 90°-rotation of the original closed shape. Thus, there was some transfer to all four shapes. Performance on the original shapes moved at 90° to their original directions of motion was not as good as performance on the originals moved in the same direction as in training (Table I, *O* and *C*); the *S*s had learned something, therefore, about the original direction of motion of the shapes. They had not, however, learned the direction of motion in relation to the shape itself, otherwise performance would have improved when 90°-rotations of the original shapes were shown moving at 90° to the original direction of motion.

Discussion: (1) Discrimination of stationary shapes. This experiment establishes for the first time that octopuses can learn to discriminate accurately between stationary shapes.[13] It may be argued that there always was some movement present in the animal's environment, even though the shapes themselves were stationary. Attacks were certainly not produced because of the movement of the opaque screen which was lifted to expose the shape at the beginning of each trial: this is proved by the long latencies of Group S, many attacks being made 20 sec. or more after the screen had been removed. It is impossible to eliminate all movement within the tank since *S*'s own breathing and locomotor movements may cause movements of the water and of any particles in the water. Throughout most of the experiment, a jet of circulating sea water was playing on the surface of the water in each tank; but, when these jets were turned off for two days toward the end of training, performance was not affected. It is true, of course, that the shapes may have moved in relation to *S*'s retina as *S* itself moved. There is no reason, however, to suppose that this played any part in the discrimination, because well-trained *S*s often remained completely at rest when the negative shape was presented.

---

[12] N. S. Sutherland and A. E. Carr, The visual discrimination of shape by *Octopus:* The effects of stimulus size, *Quar. J. exp. Psychol.*, (in press).

[13] The eye of the octopus forms a single image. K. H. Heil (Beiträge zur Physiologie und Psychologie der Springspinnen, *Z. vergl. Physiol.*, 23, 1936, 1-25) has found visual recognition of immobile prey in the jumping spider, another invertebrate with the same type of eye.

Given that octopuses can discriminate stationary shapes, the question arises of whether they do so by analyzing the same properties of shapes as when the shapes are in motion. The fact that there was a difference in performance between the subgroups H and V when the animals were trained with stationary shapes, whereas no such differences occurred between the corresponding subgroups trained with moving shapes, suggests that moving and stationary shapes may not be discriminated in terms of the same properties. One reason has already been given above, however, for supposing that this result is due to experimental error. Two further reasons may be added here: (a) In a more recent experiment with stationary shapes, no difference in discriminability was found between shapes differing only in their horizontal projections (H-shapes) and shapes differing only in their vertical projections (V-shapes). (b) If octopuses trained on stationary shapes rely more heavily on differences in the horizontal and vertical projections than do animals trained with moving shapes, then Group S should show less transfer to rotated shapes than Group M, but in fact they transferred slightly better (*cf.* shapes No. 21 and 27, Columns 6 and 7 in Table I). If, then, we set the difference in training performance between Groups M and S down to experimental error, we must consider whether the results of the transfer-tests are the same for the two groups. The correlation between rankings of transfer-shapes obtained by Groups M and S ($\rho = 0.69$) is practically as high as that between subgroups V and H, and is higher than any of the coefficients calculated for pairs of subgroups of four *S*s each. Moreover, if we consider the few shapes for which there is a wide difference between the ranks given by Groups M and S (Nos. 4, 6, 9, 12, 13, 18, 21), no consistent pattern of differences emerges. The consistency in the rankings of transfer-shapes yielded by the *S*s trained with stationary and moving shapes makes it safe to assert that both groups were analyzing the training-shapes in terms of the same—or almost the same—properties.

(2) The dimensions along which shapes are analyzed. One purpose of the present experiment was to establish the dimensions in terms of which the training-shapes used were discriminated by octopuses. The problem is to find a dimension along which the shapes can be ordered which will

account for the rank-order found in the transfer-tests. Several such dimensions have been considered, but none was found which would fit the data in a completely satisfactory way. To illustrate some of the problems of interpreting the present results, two possible ways in which animals might analyze shape will be discussed.

(a) Ratio of contour to area. On the basis of transfer-tests made in an earlier experiment, it was concluded that the members of the first two pairs of shapes shown in Fig. 1 were being distinguished in terms of the ratio $C/A^{1/2}$.[14] One purpose of the present experiment was to test this hypothesis by using training-shapes with different ratios (6.5 and 4.9 as against 6.0 and 3.8 in the previous experiment) and discovering whether this shifted the point at which shapes ceased to be treated like the original closed shape. No such shift occurred. The ratio $C/A^{1/2}$ for all shapes used in the present experiment is given in Column 10 of Table I. Since the ratios of the three rectangles (5.4) and of the triangle (4.6) are nearer to that of the original closed shape than to that of the original open, the hypothesis predicts transfer to these four shapes from the original closed shape. Instead, there was marked transfer to all four from the original open shape, as in the previous experiment (Nos. 4, 5, 7, 16).

Moreover, if shapes are ordered in terms of the ratio $C/A^{1/2}$ (omitting the two black shapes) and their ranks on this order compared with the final ranks given by the transfer-test results (Columns 1 and 11 in Table I), the correlation-coefficient is only 0.51. Although in the right direction, the correlation is comparatively low, and it is most unlikely, therefore, that the ratio $C/A^{1/2}$ was the main dimension being analyzed.

(b) Vertical and horizontal projections. If shapes were being analyzed in terms of these projections,[15] then the *S*s trained with H-shapes should have performed better in training than those trained with V-shapes, since previous evidence indicates that the horizontal projections determine behavior more readily than the vertical projections.[16] The prediction was not confirmed for Group M. Although Group S-H did perform significantly better than Group S-V, the result probably was due to experimental error.

A second way of testing whether the projective theory is correct is to examine the results of transfer-tests. We should expect in terms of the model that Groups H and V would show different amounts of transfer to some shapes. The relative amount of transfer for the two groups should depend upon the orientation of the transfer-shape: the *S*s in Group H should show transfer from the original open to further shapes which resemble it on the *horizontal* projection; while the *S*s in Group V should show transfer from the original open to further shapes which resemble it on the *vertical* projection. Inspection of Columns 8 and 9 of Table I suggests that this does occur with some shapes: thus, the upright 'T' (No. 2) is treated as a more open shape by Group H than by Group V—because of its long horizontal bar, its horizontal projection resembles that of the original open

---

[14] Sutherland, *op. cit., J. comp. physiol. Psychol.,* 53, 1960, 104-12.

[15] Sutherland, *op. cit., Nature,* 179, 1957, 11-13; *op. cit., Nature,* 186, 1960, 840-844.

[16] Sutherland, *op. cit., Brit. J. Psychol.,* 51, 1960, 9-18; *op. cit., J. comp. physiol. Psychol.,* 54, 1961, 43-48.

H-shape more than its vertical projection resembles that of the original open V-shape.

By similar reasoning, we should expect the horizontal rectangle and the 'M' to resemble the original open shape more for Group H than for Group V, and the difference should be reversed for the vertical rectangle, and the 'U' on its side. All these expectations are fulfilled (see Columns 8 and 9 of Table I for shape, Nos. 4, 20, 16, 10). There is little difference in the treatment of two shapes by Groups H and V where a difference would have been expected (Nos. 2, 13). The fact that transfer to the rotated open shape (No. 21) is less good than to many other open forms is surprising and important; exactly the same result was obtained in the earlier experiment.[17] This finding also can be explained on the hypothesis that the shapes were being analyzed along their horizontal and vertical projections; when the *open* shape is rotated it produces a very similar pattern to that produced by the original *closed* shape on the projection which differentiates the original training-shapes, *i.e.* on the horizontal projection for the *S*s trained with horizontal shapes, and on the vertical projection for those trained with vertical shapes (Fig. 1).

Most, but not all, of the remaining results of transfer-tests can be understood in terms of this theory. Shapes yielding strong peaks on one or the other of the two projections tend to be treated as open shapes, *e.g.* the horizontal and vertical rectangle (Nos. 4, 16), the triangle (No. 7), and the outline of the original closed shape (No. 14). An upright cross (No. 11) is treated as more open than 'X' (No. 17). It is interesting that, whereas a filled-in square (No. 27) is treated as more closed than a filled-in diamond (No. 26), an outline-square (No. 14) is treated as more open than an outline-diamond (No. 19); this result would be expected because, although the filled-in square yields less of a peak than the filled-in diamond, the outline-square yields stronger peaks on both projections than the outline-diamond.

Two results cannot be explained in terms of an analysis of horizontal and vertical projections: the oblique rectangle (No. 5) yields no peak on either projection, yet it is treated as an open figure; the circle yields more of a peak than the square on both projections, yet it is treated as a less open figure.

It is possible to test against the transfer-data various other hypotheses about the properties in terms of which octopuses analyze shape. Several of these have been considered (*e.g.* a suggestion put forward by Deutsch[18]), but none has been found to yield as good a fit as the second hypothesis considered above. Although the writer has not been able to formulate an hypothesis that will account for all the results reported here, the data do present a challenge to explanation. Considering the horizontal and vertical projections of the shapes makes some puzzling results fall into place—for example, it explains the (intuitively) surprising finding that transfer from the original open shape to very different open shapes (such as a cross) is actually better than it is to the 90° rotation of the open shape itself. The data are interesting precisely because many of the results are surprising to common sense, and because, in four other experiments in which animals have been trained to discriminate open and closed shapes different from those used here, a strikingly similar ordering of transfer-shapes was found. The similarity between the rankings of transfer-shapes obtained when different training figures are used suggests that the octopus may be analyzing shapes along a small number of dimensions, and this means that with sufficient ingenuity it should be possible to identify these dimensions.

---

[17] Sutherland, *op. cit., J. comp. physiol. Psychol.*, 53, 1960, 104-112.

[18] J. A. Deutsch, Theories of shape discrimination in *Octopus, Nature,* 185, 1960, 443-445.

SUMMARY

Sixteen octopuses were trained to discriminate between open and closed shapes. Half the *S*s were trained with moving shapes and half with stationary shapes. After training, the *S*s were tested for transfer to 30 different shapes. It was shown for the first time that octopuses are fully capable of discriminating stationary shapes. Provided that care is taken to maintain their readiness to attack, they do so about as readily as they discriminate moving shapes. Moreover, the *S*s trained with stationary and moving shapes transferred to other shapes in very similar ways and this suggests that they had learned the original discrimination by analyzing the same properties of the shapes irrespective of whether they were stationary or in motion.

The shapes with which the *S*s were trained and the transfer-shapes presented were selected in such a way as to test certain hypotheses about the dimensions along which octopuses analyze shapes. The following points emerged: (1) When moving shapes are used, octopuses do not discriminate more readily between shapes differing in their horizontal projections than between shapes differing in their vertical projections. (2) The results of transfer-tests showed that the *S*s did not discriminate the training shapes in terms of the ratio $C/A^{\frac{1}{2}}$ as was suggested by the results of a previous experiment on open and closed forms. (3) Many, but not all, of the results of the transfer-tests can be explained on the assumption that animals are analyzing shapes in terms of their horizontal and vertical projections. Thus, even when the training-shapes used differ in degree of closure, octopuses seem to analyze them in terms of some horizontal and vertical coördinate-system. (4) Although many of the results of transfer-tests are surprising to common sense, the same results can be replicated in other experiments using open and closed training shapes different from those used here. This suggests that octopuses analyze shapes along a limited number of dimensions.

MODIFICATION OF STEREOSCOPIC DEPTH-PERCEPTION

By HANS WALLACH, Swarthmore College, MARY E. MOORE,
University of Pennsylvania, and LINDA DAVIDSON,
Swarthmore College

This is the first in a series of studies which explore artificial conditions that yield rapid perceptual learning and the nature of the effects obtained in learning. The following strategy for discovering conditions for rapid perceptual learning was employed: We selected two cues of different types which pertained to one and the same perceptual property, such as distance or depth; by means of an optical or mechanical arrangement one of the cues was so altered that the objective situation was misrepresented; and then *O* was exposed simultaneously to the two cues, which were now contradictory.

Method. The effect to be reported here involved the perceived depth of small objects. The cues which were opposed were retinal disparity and the conditions of stimulation that gave rise to the kinetic depth-effect (*KDE*). The latter consist in changes in the retinal image of a three-dimensional object which, due to its rotation or the locomotion of *O*, alters its orientation relative to the eye; such changes are good cues for the object's shape and produce veridical three-dimensional perceptions of form.[1] For a given object, the two types of cue would, of course, tend under ordinary conditions to produce the same three-dimensional shape. We brought them into conflict by using a mirror-arrangement that increased all retinal disparities inherent in the two retinal images of an object. Such an arrangement is called a telestereoscope, and it enhances disparities by increasing, in effect, the interocular distance (Fig. 1).

Various three-dimensional wire forms, which were made to rotate about a vertical axis, were observed through the telestereoscope. Under such conditions, the wire form appears to distort as it rotates. An imaginary line between two points on the object that lie in the same horizontal plane changes in length continuously during rotation; it appears shortest when it lies in the frontal-parallel plane and longest when it points away from the observer. When it lies in the frontal-parallel plane, its length is determined by the distance between the retinal images of these points; when it occupies the depth-dimension, its apparent length depends in addition on the disparity of its end-points and suffers the telestereoscopic enhancement. Thus, each cross-section of the wire form changes periodically in width; it is narrowest when it lies

---

* Received for prior publication February 17, 1963. This work was originally supported by the U. S. Naval Training Device Center under Contract N61339-511 and was continued under a grant from the National Science Foundation.

[1] Hans Wallach and D. N. O'Connell, The kinetic depth effect, *J. exp. Psychol.*, 45, 1953, 205-217.

in the frontal-parallel plane and widest when it lies in the sagittal direction, and different cross-sections go through different phases of expansion and contraction simultaneously, depending on their momentary orientation to the sagittal direction.

Viewing the rotating wire form through the telestereoscope for 10 min. changes the relation between disparity and perceived depth. This can be demonstrated in two ways. After the 10 min. 'training period' the wire form has smaller apparent depth, and this is ascertained by having *O* estimate the perceived depth of the stationary wire form before and after the training period. The other demonstration consists in having *O*, after the training period, look directly at the rotating wire form. He will now see it distort, although ordinarily the directly viewed rotating form appears rigid. The distortion is the opposite to that observed through the telestereoscope; cross-sections will appear to decrease in width as they turn from the frontal-parallel to the sagittal orientation, indicating that the now normal disparity gives rise to a reduced depth. Although this latter demonstration of the training effect is simple and impressive, we used the estimation-procedure in most of our experiments because it yielded quantitative results. Unless otherwise reported, our telestereoscope increased the effective interocular distance in all experiments by 7.6 cm.; since the average interocular distance amounts to 6.5 cm., this more than doubled normal disparity. For each experiment, the telestereoscope was set in such a way that *O* viewed the wire form through the mirrors with the same convergence as he would view it directly. This produced a discrepancy between convergence-distance and accommoda-

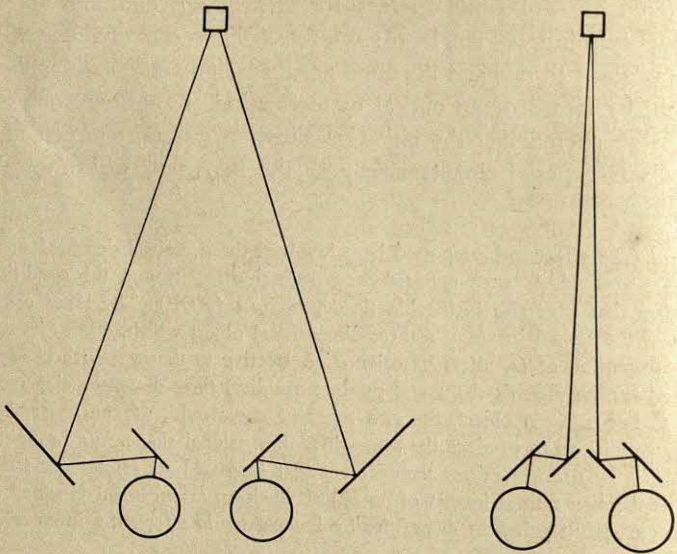


FIG. 1. ARRANGEMENTS IN ENHANCING (LEFT) AND REDUCING (RIGHT) RETINAL DISPARITY

The light-paths from the target (square) to the eyes (circles) through the four mirrors of the telestereoscope are traced.

tion-distance of 6.4 cm.,[2] which is equivalent to a negligible 0.15 diopters even for the shortest of the viewing distances used in our experiments.

BASIC EXPERIMENTS

A typical experiment employed a black wire cube with a 4.9-cm. edge, with one diagonal axis forming the extension of the vertical supporting shaft, about which the wire form could be rotated. The center of the cube was on a level with the *O*'s eyes and 63.5 cm. (25 in.) from the mirrors. Sixteen experimentally naïve *O*s, previously tested for adequate stereoscopic vision, participated. The experiment began with *O* viewing, through the telestereoscope, the stationary cube which was so oriented that one of its diagonals fell into *O*'s sagittal plane. *O* was instructed to adjust the length of a metal rod to equal the apparent height of the cube, along the diagonal. In making this adjustment, *O* looked alternately through the telestereoscope at the cube and downward at the metal rod, which he held in his hands. Then he made an estimate of the depth of the cube by adjusting the length of the metal rod to the apparent length of the sagittal diagonal. When these measurements were completed, *O* looked through the telestereoscope and watched the cube rotate at a speed of 12 r.p.m. for 10 min. At the end of the training period, *O* was asked to close his eyes while the rotation was stopped, and the cube was turned by hand into the same position it had before rotation started. *O* made another estimate of the cube's depth, and one of its height, in that order. The mean depth-estimate before rotation was 9.96 cm., while that for the height amounted to 5.71 cm. Since, objectively, the height and depth of the wire form were the same, the difference between measured depth and height (amounting to 4.25 cm.) is, of course, due to the telestereoscopic effect. After 10 min. of watching the cube rotate the mean depth-estimate was 8.36 cm., an average reduction of 1.60 cm. in perceived depth, a difference significant beyond the 0.1% level.

Similar results were obtained with a different wire form, a slightly irregular pyramid with a four-cornered base. (The edges of the base ranged from 5.1–6.3 cm. in length, and the four rods, rising to the apex, from 5.6–5.9 cm.) The perpendicular distance from apex to base was 3.8 cm., and, since estimates were made with the base in a frontal-parallel position, this distance also was the measured depth of this wire form. The rotating shaft was attached at right angles to one of the edges of the base. The mean estimate of depth (20 *O*s) was 6.93 cm. before the training period and 5.53 cm. afterwards. The mean change in depth, 1.40 cm., was significant beyond the 0.1% level. In this experiment, the diagonal of the pyramid's base also was measured before and after the training period. Since the diagonal lies in the frontal-parallel plane when measurements are taken, a change in its length reflects a change in perceived size. Mean estimates of the length of the diagonal were exactly the same before and after the training period, indicating that no change in apparent size accompanied the large change in perceived depth. In nearly all experiments to be reported, estimates of frontal-parallel size were made before and after the training period to assure that the changes in the depth-estimates obtained were not related to changes in over-all size.

---

[2] This is the increase in the true viewing distance due to the detour of the light-path via the outer mirrors, which concerns accommodation but not convergence.

ALTERNATIVE EXPLANATIONS

There are two other interpretations of the effect of the training period besides perceptual learning which need be discussed.

(a) One alternative is that the period of rotation reveals the true shape of the wire form and the final estimate of depth is influenced by the memory of this shape. This explanation does not apply to the qualitative demonstration of the effect of training (the distorting of the wire form in normal vision following training), but should be considered with regard to our quantitative method. It is easily refuted by the fact that the training effect can be transferred to a different wire form.

In these experiments, the initial and final estimates of depth were made with one wire form while a different one was used in training. Except for this exchange of the objects presented, the procedure here was the same as in the basic experiments. Table I lists the magnitude of five effects of transfer and permits comparison with the results obtained when test- and training objects were the same. No conclusions should be drawn from apparent differences among the various effects of transfer; they do not approach significance. Most of the data were obtained early in our study before improved techniques were introduced; only Experiments *A* and *B* were done with groups of *O*s of whom none had previously served in telestereoscopic experiments and under conditions that afforded cues of perspective distance (see below). There can be no doubt that the training effect transfers to different objects, but it seems to do so with some loss.

(b) The second alternative interpretation of our effect is that the diminished depth after the training period is an effect of satiation. This term refers to the fact that prolonged exposure to the same pattern or

the same continuous movement is apt to lead to a deterioration of the corresponding perceptual experience. Such spontaneous changes can be observed when an outline-figure is fixated and becomes slightly smaller, or when stroboscopic motion between two spots which light up alternately is viewed for a prolonged period, conditions which lead eventually to cessation of the apparent motion. Diminished depth might conceivably be the result of a similar change.

Experiments with diminished interocular distance. We tested this explanation with experiments which involved the reverse training. The mirrors of the telestereoscope were so rearranged that, rather than increasing disparity, they reduced it (Fig. 1) and therefore diminished perceived depth in relation to objective depth. Training with this device should lead to increased stereoscopic depth. If here, too, the training period should turn out to reduce perceived depth, an interpretation in terms of satiation would be indicated, whereas an increase in depth measured by the final estimate would point to learning as the cause of our effect. The new device reduced the effective interocular distance by 2.8 cm. Irregular pyramids, similar to the one used in the basic experiment, were the wire forms used. The base remained constant in all pyramids, but, since the new arrangement of the mirror reduced depth, the pyramids used in the present experiment had greater distances between apex and base; namely, 6.4 and 12.2 cm. For the former (25 in. distance from eyes, 17 *O*s) the mean depth-estimate before the training period was 3.03 cm.; after rotation it was 3.59 cm., the increase in depth of 0.56 cm. was reliable beyond the 0.5% level. For the latter (30 in. from eyes, 16 *O*s) the depth-estimate

TABLE I

THE RESULTS OF EIGHT EXPERIMENTS COMPARED

The diagonal cube (*A–D*, *G*, *H*) and the pyramid (*E*, *F*) have been described under Basic Experiments. The elongated pyramid (*G*) had a base identical with that of the pyramid, but a distance from apex to base of 6.4 cm. The perpendicular cube was in normal orientation and had an edge of 5.3 cm. The octohedron was regular and its edges were 6.3 cm. long. All decreases in depth were significant beyond the 1% level.

	Test object	Training object	*N*	Depth-decrease (cm.)
A	Diagonal cube	Diagonal cube	13	1.62
B	Diagonal cube	Perpendicular cube	18	1.24
C	Diagonal cube	Diagonal cube	16	1.60
D	Diagonal cube	Perpendicular cube	24	0.83
E	Pyramid	Pryamid	20	1.40
F	Pyramid	Diagonal cube	16	0.59
G	Diagonal cube	Elongated pyramid	14	0.73
H	Diagonal cube	Octohedron	21	0.68

changed from 5.71 to 6.57; this difference of 0.86 cm. was significant beyond the 0.1% level. The increase in estimated depth due to the training amounted to 28% and 22% of the presumed decrement in depth caused by the mirror-arrangement. These relatively low values were probably due to the relatively small change in the effective interocular distance which the present mirror-arrangement produced. The telestereoscope used in the basic experiment augmented effective interocular distance by 7.6 cm., an increase of 117% (when the natural interocular distance is assumed to be 6.5 cm.), while the present mirror-arrangement reduced it by 2.8 cm., causing a comparable percentage change of only 74%. These considerations, together with the high reliability of the results obtained with fairly small groups of *O*s, make it clear that deterioration of depth-perception need not be seriously considered as an explanation for the major portion of our effect.

DISSIPATION OF THE EFFECT OF TRAINING

Once the new evaluation of disparity has become established through training, it should be possible to unlearn it—to relearn the former evaluation—by watching an object rotate in normal vision, without the telestereoscope. Since, after training, the object in normal vision is seen with reduced depth, stereoscopic depth-perception is in disagreement with the *KDE,* and the object seems to distort. The procedure we used was to have *O* look through the telestereoscope at the pyramid, make a height- and a depth-estimate, watch the pyramid rotate for 10 min., and then, before making the final depth-judgment through the telestereoscope, watch another pyramid in rotation for a brief period, without the telestereoscope. This second pyramid had more depth than the original one in order to make it similar to the original pyramid as seen through the telestereoscope; the distance from apex to base was 6.4 cm. Table II shows the decreases in depth measured by this procedure with various durations of the un-

TABLE II

UNLEARNING

N	Unlearning period (min.)	Depth decrease (cm.)	Reduction in learning effect (%)
20	0	1.40	—
18	0.5	1.08	23
12	2.0	0.58	59*
15	5.0	0.39	72*

* Significant beyond the 1% level.

learning period, as well as the difference between the depth-decrease in the learning experiment and the depth-decrease measured after each of the unlearning periods. Unlearning operated as expected, counteracting part of the depth-decrease that resulted from the original training, but, even after 5 min. of unlearning, the learning effect that remained was significantly different from zero beyond the 2% level.

More interesting than dissipation through unlearning is the dissipation of the training effect through mere lapse of time. If this dissipation were very rapid, it could be a source of considerable variance in our measurements of the training effect, for the time taken to make a depth-estimate varies from *O* to *O*. Information on the loss of the training effect with passing of time is, therefore, of practical concern in our experiments. It is also of theoretical interest; after having been modified, does the stereoscopic perception of depth return to normal only because it is again practiced under normal conditions, or is mere passage of time sufficient to reëstablish the normal relation between disparity and perceived depth?

We tried to answer this question by delaying the second estimate of depth. At the end of the training period, *O* sat quietly with eyes closed for either 5 or 12 min. before making a second estimate. The mean depth-changes obtained for the pyramid after these forgetting periods are given in Table III. It might be added that the effect of 0.38 cm. remaining after the 12-min. intermission still is significant beyond the 0.5% level. The experiment with the 5-min. 'forgetting' period was recently repeated under improved viewing conditions (see below) with the same wire form. With the second estimate made immediately after training, a group of 20 *O*s gave a mean depth-decrement of 1.28 cm., while from another group of 20 *O*s, who gave their second estimates with 5-min. delay, a mean decrement of 0.81 cm. was obtained; the difference of 0.47 cm. between the groups was significant beyond the 2% level.

TABLE III

FORGETTING

N	Forgetting period (min.)	Depth-decrease (cm.)	Reduction in learning effect (%)
20	0	1.40	—
10	5	0.89	36
15	12	0.38	73*

* Significant beyond the 5% level.

Whether one wants to call the dissipation of the training effect with lapse of time 'forgetting' depends on what view of the forgetting process one favors. If one assumes that forgetting usually consists in a replacement of a newly acquired habit by older habits, the term can be applied here. The discovery of a rapid modification of the relation between disparity and perceived depth makes it very likely that the original relation which *O* brings to our experiments also is an acquired one and is temporarily superseded by our training effect.

CUMULATIVE EFFECT OF TRAINING

We consider our modification of stereoscopic depth an instance of learning. Perhaps the most important question to be asked in this context is whether it is a continuous process, that is, whether our training effect increases with the length of the training period. We did an experiment in which the training period was twice interrupted to obtain depth-estimates for shorter training times. The first post-training depth estimate was made after only 1 min. of observing the rotating wire pyramid through the telestereoscope. Immediately thereafter—no height estimate was made—observation of the rotating wire form was resumed and lasted for 3 min. A new depth-estimate was made, followed by another training period lasting 6 min., after which a final depth-estimate was obtained. The results are given on the left side in Table IV. They show that virtually the whole effect was produced in the first two training periods; the last period of 6 min. did not seem to add to our effect.

This seemed an unlikely result and we were unwilling to accept it as genuine, and this reluctance, in turn, led to an important improvement of our experimental conditions. This improvement has to do with the fact that perceived stereoscopic depth depends not only on the given disparity, but also on the distance of the object that causes the disparity; the greater the distance, the greater will be the apparent depth that results from the given disparity.[3] This means that a change in apparent depth of a wire form could come about through a change in the evaluation of a cue for distance. Our training effect is protected against such an interpretation by the finding that apparent size is not altered by the training period. If the training effect were produced by a change in the evaluation of distance cues, a size- as well as a depth-change should result. Changes in the *effectiveness* of available distance cues with prolonged observation could, however, account for our unlikely result. With two exceptions, all experiments thus far reported were done under conditions in which only convergence and accommodation operated as cues for distance; it seemed just possible that this kind of cue deteriorates with prolonged observation.

---

[3] J. Linschoten, *Structuranalyse der binocularen Tiefenwahrnehmung: Eine experimentelle Untersuchung,* 1956, 180.

To strengthen the representation of the distance of the wire form, we introduced a further distance-cue—perspective—to the viewing condition. A wooden platform provided a horizontal surface stretching from *O* to the shaft supporting the wire form and 3 ft. beyond. This surface was only 16 cm. below the center of the wire form and started being visible approximately 15 in. from *O*. To strengthen its perspective effect, we covered it with a checkered oilcloth. The checkered surface was visible only to one eye in order to prevent telestereoscopic depth-perception of the pattern. A low screen in front of the left-objective mirror blocked the view of the horizontal surface for the left eyes, while the wire form was visible to both eyes.

TABLE IV

Successive Determinations under Two Conditions

	Without perspective-cues ($N=24$)			With perspective-cues ($N=41$)		
	Mean depth (cm.)	Mean difference	*t*	Mean depth (cm.)	Mean difference	*t*
Depth prior to rotation	6.67			7.17		
		0.47	2.24†		0.37	2.46*
Depth after 1 min. rotation	6.20			6.85		
		0.45	3.46*		0.30	2.73*
Depth after total of 4 min. rotation	5.75			6.55		
		0.06	0.46		0.26	3.71*
Depth after total of 10 min. rotation	5.70			6.29		

* Significant beyond the 1% level. † Significant beyond the 5% level.

When the experiment was repeated under these conditions, each of the three training periods produced a clear and statistically significant effect (right side of Table IV). When the three training effects were plotted against a logarithmic time-scale, an almost straight line was obtained (see center curve in Fig. 2). An analysis of variance showed significance beyond the 0.1% level for the total training effect ($F = 12.21$ with 3 and 120 *df*). For further comparisons, the assumption was made that the time-intervals between the four depth-estimates were 'psychologically equal.' When the mean depth-estimates were plotted against such a scale, it was found that again they nearly fell on a straight line. A linear trend was fitted to the group means, and the mean slope of the group curve was then tested against the variability among the slopes of the curves for the individual *O*s.[4] The linear trend was found to be significant beyond the 0.1% level. ($F = 14.97$ with 1 and 40 *df*). This indicates that the data for the individual *O*s as well as the group-data can be described as a linear function, and shows that our training effect is cumulative.

In addition to being highly significant, the linear component of the training effect accounted for virtually all the variance due to the training conditions (99.87%). The minute residual (0.13) eliminates the possibility of any other systematic source of variance. This spoke well for the improved conditions under which the experiment was done. All further experiments were therefore done with the checkered surface visible, and the more important earlier experiments were repeated under these conditions.

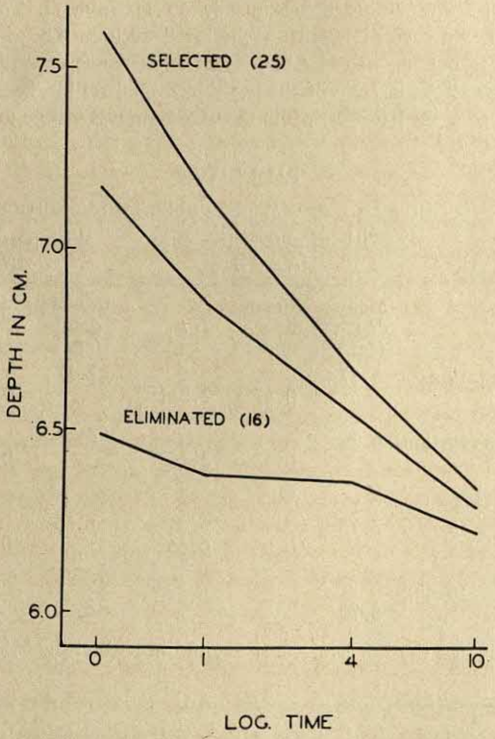


FIG. 2. CHANGES IN DEPTH AS A FUNCTION OF TIME

---

[4] D. A. Grant, Analysis of variance tests in the analysis and comparison of curves, *Psychol. Bull.*, 53, 1956, 141-154.

Among them was one that had produced a surprising result. In addition to the two series of experiments on the dissipation of the training effect just reported, a third one had been done where *O*, during the interval between the end of the training period and the final depth estimate, looked steadily through the telestereoscope at the now stationary wire form. This condition apparently yielded a very fast dissipation of the training effect, as fast as the 'unlearning' condition. This inexplicable result disappeared when the experiment was repeated under the improved viewing conditions; looking at the stationary wire form did not cause the training effect to dissipate faster than mere lapse of time spent with eyes closed. It looks as if prolonged inspection of the wire form with only oculomotor cues for distance present produced a slight effect of increasing depth which is superimposed on our training effect.

An inspection of the individual records that contributed to the right side of Table IV made it appear that those *O*s who showed a larger initial depth also showed a larger training effect. We therefore divided the 41 *O*s who had participated in the experiment into two groups according to their initial depth and computed the mean effects for the three training periods for each group. A mean of 7.17 cm. had been obtained for the initial depth for all *O*s, and their mean height, a measure of frontal size, was 6.76 cm. We grouped *O*s according to whether their initial depth-estimates were higher or lower than their height-estimates. The groups thus constituted consisted of 25 *O*s with a mean initial depth of 7.60 cm. (selected *O*s) and 16 with a mean initial depth of 6.49 cm. (eliminated *O*s). The upper and the lower curves in Fig. 2 show the mean training effects of the two groups plotted against an abscissa of log time in sec. plus one. The curve for the eliminated *O*s shows a very small training effect indeed—only 0.27 cm. for the three training periods combined—while the corresponding figure for the selected *O*s is 1.26 cm. Just why *O*s whose initial depth estimates were lower showed much less of a training effect is still a matter for speculation; but selection of *O*s according to their initial depth-estimates can be used to keep variance due to individual differences low, which becomes important where one tests for a difference in the amount of the training effect between different groups of *O*s.

We return now to the main result of the experiment, that our effect grows nearly linearly with the logarithm of the training time. This means, of course, that increments in the effect per unit of real time become smaller with elapsed total training time. If this relationship prevails for longer training times, as appears likely, little gain can be expected from prolonging the training period beyond 10 min. To be sure, the additional training effect produced during the last period seems somewhat more reliable ($t = 3.71$) than the one for the first period ($t = 2.46$), but this advantage appears small when weighed against increased fatigue and boredom of *O*. Thus far we have not worked with a training period longer than 10 min. A truly linear function of log time would predict that it would take 2 hr. to double the training effect that is obtained in 10 min. and that total elimination of the effect of the telestereoscope would be a matter of many days; but with a lifetime of practice in the normal relation between disparity and perceived depth opposing the training effect, it seems questionable that the latter could ever happen. It is also conceivable that stereoscopic depth-perception is not the only function which is changed in the training period; it is possible that the *KDE*, too, undergoes an alteration that serves to diminish the discrepancy which the telestereoscope produces. If this were the case, the modification of stereoscopic depth perception would be expected to halt before it fully compensated for the enhancement of disparity caused by the telestereoscope. The next experiment to be reported takes up this issue.

MODIFICATION OF THE KINETIC DEPTH-EFFECT

Just as it is possible to alter the relation between disparity and perceived depth, it is conceivable that the relation between the deformation of the image of a turning object and its perceived depth can be modified. This might possibly happen in our experiments, provided it is the contradiction between two sets of cues which is the cause of the training effect. According to this interpretation, learning diminishes the discrepancy between the perceptual result of the kinetic depth-effect (*KDE*) on the one hand and the artificially enhanced stereoscopic depth on the other. This diminution could be accomplished in two ways, merely by the change in the relation between disparity and perceived depth which was actually obtained, or by a simultaneous modification of the *KDE*. If the *KDE* were so changed that a given image-deformation gave rise to greater depth of the perceived object, this would serve also to diminish the discrepancy between the results of the two processes of depth-perception. The fact that, in our experiments, the *KDE* yields at the outset veridical depth, and stereoscopic vision does not, is irrelevant as long as there is no further kind of depth-cue in operation which provides independent information on the true depth. We attempted in the following manner to determine whether the *KDE* undergoes such a modification in our experiments.

The wire form used was the pyramid which, because of its irregular shape, does not yield perspective depth-cues. The training lasted an uninterrupted 10 min., and consisted in the usual binocular observation of the rotating wire form through the telestereoscope. It was in the depth-estimates that the present experiment differed. They were made while the wire form was in rotation and was monocularly observed.

The results for this experiment, in which 28 *O*s participated, did not show an adaptive modification of the *KDE*. Such an effect would have resulted in an increased depth of the monocularly perceived wire form after training. In fact, mean estimated depth decreased slightly, from 4.66 cm. to 4.44 cm., and this difference was reliable at the 5% level. This decrease turned out to be a matter of a small decrement in total size as indicated by a change in the height-estimate from 6.52 cm. to 6.16 cm. When the depth-change was compared with the height-change no significant difference was found ($t = 0.52$). Inasmuch as observation was monocular when the estimates were made, a change in total size after prolonged observation is not a matter for concern. It is most likely due to the absence of convergence as a distance-cue. The experiment makes it clear that our training period does not produce a modification of the *KDE*.

MODIFICATION OF STEREOSCOPIC DEPTH BY MEANS OF PERSPECTIVE

There is a way to achieve a change in the relation between the given disparity and perceived depth which does not involve the *KDE*. In the case of regular orthogonal wire forms, stationary exposure during the training period also leads to such an effect. We demonstrated this with two such wire forms, a cube in normal orientation and a regular octahedron. They were attached to shorter shafts (the cube by means of a bar across the botton face) in such a way that *O* saw them obliquely from above. In both forms, wire squares in horizontal orientation were visible, which might by the shape of their retinal projections give away the object's true depth. That this was the case could be inferred from the fact that many *O*s' initial depth-estimates failed to show the enhanced depth that the telestereoscope usually brings about. This same fact causes difficulties for demonstrating the effectiveness of perspective cues in modifying stereoscopic depth. The *O*s whose initial depth-estimates did not show the telestereoscopic effect could not be expected to produce evidence of a training effect. Yet, being presented with two sets of conflicting cues, they might well develop a modified stereoscopic depth perception.

We circumvented this difficulty by casting our demonstration in the form of a transfer experiment. Each of the wire forms mentioned was presented in a stationary position for 10 min. viewing through the telestereoscope, but the depth estimates before and after this inspection period were made with the diagonal cube. The experiments were done before perspective cues for distance were introduced. Inspection of the stationary octahedron caused an average reduction of 0.41 cm. in the apparent depth of the diagonal cube (18 *O*s); for the cube in normal orientation, the mean difference in estimated depth before and after the inspection-period amounted to 0.47 cm. (20 *O*s). Both differences were significant at the 1% level. This effect has not yet been further investigated. It is mentioned here because in certain contexts it may be necessary to control for the operation of perspective depth-cues.

SUMMARY AND CONCLUSIONS

Experiments yielding rapid perceptual learning were described. By having *O*s view rotating wire forms through a telestereoscope for periods of 10 min., a change in the relation between the given retinal disparities and perceived depth was obtained which partially compensated for the depth-enhancing effect of the telestereoscope. This change became evident when, during a subsequent direct viewing of a rotating wire form, *O*s perceived it distorting. The change could be measured by having *O*s make estimates of depth before and after the training period. By using an arrangement of mirrors that diminished disparities, an opposite training effect could be obtained, which caused given disparities to produce greater than normal depth. It was possible to present one wire form for the training period and use a different one for the estimation of depth, but in this transfer situation the training effect seemed to be diminished. The training effect dissipated quickly with mere passage of time, but not as rapidly as it was acquired. When the growth of the effect during training was studied, it was found to be cumulative, not only for a group of *O*s but also for individual *O*s.

These findings raise many problems. What is the ultimate cause of our training effect: the fact that enhanced stereoscopic depth implies a distortion of the rotating object or the contradiction between the two sets of depth cues present, disparities on the one hand and conditions of stimulation for the *KDE* on the other? Why is the training effect diminished when it is transferred to another wire form? Why is it possible that a sizeable modification of stereoscopic depth-perception can be obtained in so brief a training period? Work on these questions is in progress.

PREFERENCES OF THE WHITE RAT FOR SOLUTIONS OF SUCROSE AND QUININE HYDROCHLORIDE

By PAUL THOMAS YOUNG, R. G. BURRIGHT, and L. J. TROMATER,
University of Illinois

It was shown by Young and Christensen that hedonic processes summate algebraically when different concentrations of sucrose and sodium chloride are combined in a single compound solution.[1] The present investigation was undertaken to explore hedonic interactions in another area of gustatory stimulation. Compound solutions that contain both sucrose and quinine are ideal for the study of algebraic summation of hedonic processes because it is known that rats, when given a brief choice, prefer the higher sucrose concentration,[2] and that animals are aversive to solutions of quinine at most if not at all concentrations.

The present investigation is concerned with two main questions. First, what are the thresholds of aversion to solutions containing quinine hydrochloride when distilled water and different concentrations of sucrose serve as a base for threshold-determinations? Secondly, what concentration of quinine must be present to make a compound sucrose-quinine solution isohedonic to a standard sucrose solution of lower concentration? There is an incidental question: Are some quinine solutions preferred to distilled water?

METHOD

Apparatus. The apparatus used in this study has been described in detail elsewhere.[3] It consists of six boxes which can be used for testing six rats simultaneously. At the front of each box are two burettes with drinking spouts projecting through a lucite wall into the interior. When a rat touches its tongue to fluid at the tip of a nozzle, a circuit is closed through the floor of the box and the animal's body. A subthreshold current is amplified to operate relays and counters.

---

* Received for publication September 29, 1962. The writers are indebted to Professor William E. Kappauf for valued assistance. The work was supported by Grant No. G-14538 from the National Science Foundation.

[1] P. T. Young and K. R. Christensen, Algebraic summation of hedonic processes, *J. comp. physiol. Psychol.*, 55, 1962, 332-336.

[2] P. T. Young and J. T. Greene, Quantity of food ingested as a measure of relative acceptability, *J. comp. physiol. Psychol.*, 46, 1953, 288-294.

[3] P. T. Young and W. E. Kappauf, Apparatus and procedures for studying taste-preferences in the white rat, this JOURNAL, 75, 1962, 482-484.

Procedure. In the present investigation, the same pair of test-fluids was presented in all six boxes of the electronic tester. Six rats were placed in the testing boxes successively at 10-sec. intervals, and 10 sec. after the last animal had been deposited *E* read the pair of counters for the first box. Every 10 sec. thereafter, he read a pair of counters for successive boxes in the row, thus obtaining one pair of readings per min. for each box throughout the exposure-period. At the close of a test, the animals were removed from the boxes in the original order at 10-sec. intervals. A record thus shows the cumulative number of tongue-contacts, minute by minute, for each animal, throughout a 15-min. exposure-period. With this procedure, it was possible to test six rats simultaneously in little more time than required for one.

At the start of the investigation, we ran two tests for each pair of fluids—one test per day with burettes in each spatial order. Later we found it feasible to reverse positions between successive daily tests while steadily increasing the concentration of quinine.

Effects due to order of presentation are difficult to control in work with quinine solutions because the higher concentrations markedly inhibit the animals and may give a persistent aversion to both test fluids. After preliminary observations, we decided to work only with ascending series of quinine concentrations and not to exceed a concentration of 0.25 gm./100 ml. solution.

Solutions. Solutions were prepared and specified as weight/volume ratios. In simple solutions, the concentration is expressed as the number of grams of solute dissolved in 100 ml. of solution and in compound solutions the concentrations of two solutes, sucrose and quinine, are expressed as grams per 100 ml. of solution. In earlier work such weight/volume ratios were called *percentages* but this is a misnomer.[4]

All solutions were prepared with distilled water. For sucrose, we used a good quality of commercial cane sugar. Chemically, sucrose is $C_{12}H_{22}O_{11}$ and its molecular weight is 342.3. For simplicity, we refer to sucrose as *CHO*. Chemically pure quinine hydrochloride crystals were used. Chemically, quinine hydrochloride is $C_{22}H_{24}O_2N_2$. HCl and its molecular weight (anhydrous) is 360.88. For simplicity, we refer to quinine hydrochloride as *Q*.

---

[4] In papers published by P. T. Young, 1944-1949 inclusive, *molal* concentrations were used, *i.e.* a 1% solution was 1 gram of solute added to 100 cc. of distilled water. Subsequently, beginning with the study by Young and Greene (*op. cit.*) in 1953, and terminating with a study by Young and Christensen (*op. cit.*), solutions were prepared and specified in terms of a true percentage by weight. This method of specification was used because of an interest, at the start, in the quantities of solute and solvent ingested; and we measured intake consistently in grams. In all current studies upon preference and activity, carried out with the electronic tester, we have prepared and specified solutions as weight/volume ratios and measured intake in volumetric units. A weight/volume ratio has several advantages: (1) it is readily understood and meaningful to nonchemists; (2) it is convenient in preparation of simple and compound solutions by dilution; and (3) specifications of concentrations as weight/volume ratios can be transformed readily to molar concentrations when the molecular weight of the solute or solutes is known. On the problem of preparing and specifying solutions see: Carl Pfaffmann, P. T. Young, V. G. Dethier, C. P. Richter, and E. Stellar, The preparation of solutions for research in chemoreception and food acceptance, *J. comp. physiol. Psychol.*, 47, 1954, 93-96.

All solutions containing Q were prepared by dilution from a stock of 1 gm./100 ml. solution. In the total experiment, we used 15 concentrations of Q. The highest concentration (Q_{15}) was 0.25 gm./100 ml. solution; the next (Q_{14}) was 0.125 gm./100 ml. solution; and so on down, each concentration being one-half of the preceding concentration. The Q-scale is shown at the base of Fig. 1.

Rats and diet. The *S*s were female rats, 50-60 days old at the start of the experiment and less than 1 yr. old at the close. They were received from the Holtzman Rat Company in three shipments.

The rats were fed *ad libitum* upon an unlimited supply of Purina rodent cubes and tap water. There was no dietary depletion, deprivation, or metabolic need at any time as far as we know. The animals were taken directly from their living cages to the testing apparatus. They were need-free. Need-free rats, of course, are not the same as satiated rats; a need-free animal may or may not be satiated upon a test-fluid such as a sucrose solution.

RESULTS

In analyzing the data, we relied upon two main methods. First, using the individual rat as a unit, we tabulated the numbers of animals within a group showing preference for one fluid and for the other. Secondly, for each individual rat we examined the absolute and relative frequencies of tongue-contacts with each test-fluid. The fluid receiving the greater number of tongue-contacts during a specified exposure-period was arbitrarily defined as the preferred fluid.

(1) *Aversion-thresholds: Quinine vs. distilled water.* In the first series of tests, 12 rats were offered a choice between distilled water and a solution of quinine hydrochloride. The series began with the lowest concentration, Q_1, and continued step by step to the highest, Q_{15}. Each quinine solution was paired with distilled water and a pair of test-fluids was presented for 30 min. with one spatial arrangement of burettes and on the following day for 30 min. with the reverse spatial arrangement. Each two-day test was scored as a unit, the fluid with the greater number of tongue-contacts being counted as the preferred fluid.

During the first six tests, Q_1 to Q_6, there was no significant preference for either fluid. At no concentration was the quinine solution preferred to distilled water. With concentrations Q_7 to Q_{15}, more rats showed a preference for water than for quinine, but the animals became increasingly inactive as the concentration rose.

Individual aversion-thresholds were determined. An aversion-threshold is defined as that concentration of quinine at which 25% of the tongue-contacts are with the quinine solution and 75% with distilled water. For purposes of the study, we discarded tests in which an individual rat made fewer than 100 contacts during a 2-day 60-min. test. With this criterion, we had to eliminate two rats which simply became inactive before meeting the criterion of aversion. The remaining ten rats showed aversion-thresholds between two quinine concentrations as follows:

Concentration of *Q*:	6		7		8		9		10		11		12		13		14		15
Number of rats:		1				2		2		2		1		1		1			

The range of these thresholds is from $Q_{6.5}$ to $Q_{13.5}$ or from 0.0007 to 0.094, and the mean is 0.0086 gm./100 ml. solution. This range and the mean are shown at the top of Fig. 1.

There was a marked tendency for the frequency of tongue-contacts to decrease with increasing concentration of *Q*. This decrease was present with *both* fluids, but from Q_7 to Q_{15} there were consistently and significantly more contacts with water than with quinine solution.

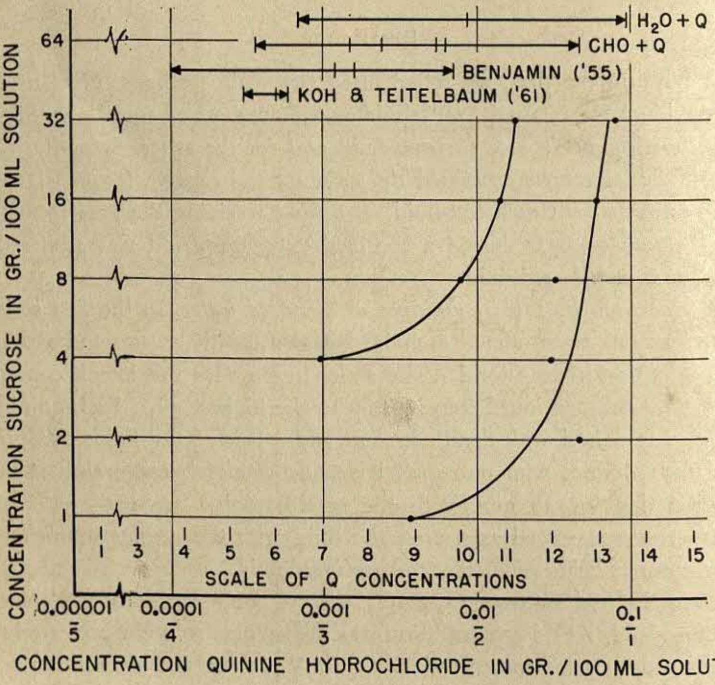


FIG. 1. AVERSION-THRESHOLDS AND ISOHEDONS IN THE SUCROSE-QUININE HYDROCHLORIDE AREA

(2) *Aversion-thresholds: Quinine vs. sucrose.* Following the above

tests with distilled water, four series of tests were made in which a simple sucrose solution was pitted against a compound solution containing the same concentration of sucrose and an ascending series of quinine concentrations. The four series were 1 *CHO* vs. 1 *CHO* + *Q*; 2 *CHO* vs. 2 *CHO* + *Q*; 4 *CHO* vs. 4 *CHO* + *Q*; 8 *CHO* vs. 8 *CHO* + *Q*. A complete preference-test required two days, since both spatial arrangements of the burettes were presented. In these and all following tests the exposure-period was reduced to 15 min./day. The first two series of tests were run with 24 rats, tested in groups of 6. Then the group was divided and the last two series were run on the same days with 12 rats in each subgroup. Each 2-day test was scored in terms of the numbers of tongue-contacts during the 30-min. exposure-period, the fluid receiving the greater number of contacts being scored as the preferred fluid.

TABLE I

NUMBER OF RATS SHOWING PREFERENCE FOR SUCROSE AND FOR SUCROSE-QUININE SOLUTIONS AT DIFFERENT LEVELS OF CONCENTRATION

N	Solutions tested	Concentration of *Q* in Compound solution														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
24	1 *CHO*						15	20	21	23	21	22	24	24	24	24
	1 *CHO*+*Q*						9	4	3	1	3	2	0	0	0	0
24	2 *CHO*	19	19	22	15	16	20	21	23	24	23	24	24	24	24	24
	2 *CHO*+*Q*	5	5	2	9	8	4	3	1	0	1	0	0	0	0	0
12	4 *CHO*	9	7	5	7	7	7	9	11	10	12	12	12	12	12	12
	4 *CHO*+*Q*	3	5	7	5	5	5	3	1	2	0	0	0	0	0	0
12	8 *CHO*	5	7	5	6	3	7	9	8	7	10	12	12	12	12	12
	8 *CHO*+*Q*	7	5	7	6	9	5	3	4	5	2	0	0	0	0	0

Table I shows the number of rats with a preference for the simple *CHO* solution and for the paired compound *CHO* + *Q*. In the first series of tests, we omitted the lower concentrations of *Q* and began with Q_6, but in the other series all 15 concentrations of *Q* were presented in ascending order.

From the data recorded in Table I, we determined aversion-thresholds for the group as a whole. An aversion-threshold for the group is defined as that concentration of *Q* at which 25% of the rats prefer *CHO* + *Q* and 75% prefer *CHO*. Aversion-thresholds were determined from group-data graphically, by linear interpolation. The values obtained are shown in Table II.

Since the group-data do not reveal individual differences, we determined an aversion-threshold for each rat in each series of tests, relying upon the relative frequencies of tongue-contacts. The individual threshold usually fell between two *Q*-values; we then placed the threshold at the mid-concentration.

Table II shows the range of individual aversion-thresholds and the mean of individual thresholds at four levels of *CHO*-concentration. The total range of individual aversion-thresholds and the four mean values are presented graphically near the top of Fig. 1. The data in Table II indicate that group-thresholds, based upon the frequency of preferences within a group, are lower at each *CHO*-level than the mean individual thresholds, based upon tongue-contacts. All aversion-threshold values, however, are of the same order of magnitude.

TABLE II

AVERSION-THRESHOLDS BASED ON DATA FOR GROUP AND FOR INDIVIDUAL RATS

Test	*N*	Group-data	Individual data	
			range	*M*
1 *CHO* vs. 1 *CHO*+*Q*	24	0.00078	0.00073–0.0470	0.0054
2 *CHO* vs. 2 *CHO*+*Q*	24	0.00036	0.00036–0.0120	0.0015
4 *CHO* vs. 4 *CHO*+*Q*	12	0.00073	0.00073–0.0059	0.0024
8 *CHO* vs. 8 *CHO*+*Q*	12	0.0052	0.00150–0.0230	0.0062

(3) *Tests with sucrose.* After the above studies of aversion-thresholds, we turned to the problem of reversing preferences by adding quinine to an an originally preferred sucrose solution. The general plan was to establish a preference between two simple sucrose solutions, and then to add quinine hydrochloride to the preferred fluid, step by step, until the preference reversed. The preliminary tests with simple sucrose solutions revealed some interesting results which will be briefly considered before we turn to the effects of adding quinine.

In the preliminary tests with sucrose solutions, all rats without exception developed a preference for the solution with the higher concentration, which result agrees with theoretical expectation.[5] The rate at which a preferential discrimination developed was a function of the difference in concentrations between the two sucrose solutions. When the difference between two concentrations was small, the animals apparently could and did learn to discriminate preferentially between the test-fluids; but, when the difference was large, the animals almost immediately showed a preference for the higher concentration.

---

[5] Young and Greene, *op. cit.*, 288-294.

In the preference test with 2 *CHO* vs. 1 *CHO*, the difference in concentration was small. Table III shows the development of a preference (referred to as a preferential trend) in this test. The table gives the number of rats, in a group of 24, with a preference for 2 *CHO* and the number with a preference for 1 *CHO*. It will be seen that repetition of this test yielded increasing frequencies of preference for the fluid with the higher concentration.

After 10 initial tests with simple *CHO*-solutions, we added the lowest concentration of quinine, Q_1, to 2 *CHO*. The test with 2 *CHO* + Q_1 vs. 1 *CHO* was made with all 24 rats; apparently the addition of Q_1 made no difference in the preference for the fluid containing 2 *CHO*. Then, with 12 of the rats, the concentration of Q was increased gradually, as shown in Table III. From Q_1 to Q_8, the addition of Q made no difference in the preference for the fluid containing the higher *CHO*-concentration. In fact, the preferential trend continued until all rats consistently preferred the fluid containing 2 *CHO*.

TABLE III

PREFERENTIAL TREND WITH TWO SUCROSE SOLUTIONS (24 RATS) AND REVERSAL OF PREFERENCE WITH THE ADDITION OF QUININE TO THE PREFERRED SOLUTION (12 RATS)

No. rats with a pref. for:	Preference test											
	2 *CHO* vs. 1 *CHO*										2 *CHO*+Q_1 vs. 1 *CHO*	
2 *CHO*	13	16	19	18	18	21	23	20	24	19	24	23
1 *CHO*	11	8	5	6	6	3	1	4	0	5	0	1

No. rats with a pref. for:	Concentration of Q in 2 *CHO*+Q vs. 1 *CHO*													
	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2 *CHO*+Q	12	11	12	12	12	12	12	11	11	9	7	5	3	0
1 *CHO*	0	1	0	0	0	0	0	1	1	3	5	7	9	12

In other tests with larger differences in sucrose concentration, the preferential trends were present but less pronounced. This fact is illustrated by results recorded in Table IV.

In the tests with 32 *CHO* vs. 1 *CHO* and 16 *CHO* vs. 1 *CHO*, a preference for the higher concentration of *CHO* appeared immediately, and repetition of these tests was not necessary; but, in the tests with 4 *CHO* as a standard, three or four daily repetitions were required to establish uniformity of preference for the sucrose solution with higher concentration.

Another interesting result came to light when we made a minute-by-minute analysis of the test with 32 *CHO* vs. 1 *CHO*. The data-sheets revealed that after the first 2 min. there was a steady decline in the drinking of 32 *CHO*. During the last 6 min., the rats were almost completely inactive, making no contact with either fluid. This pattern of behavior marks the approach to satiation upon a highly concentrated sucrose solution.

During the approach to satiation, some rats made 1 or 2 or 3 contacts with 1 *CHO* in an otherwise inactive minute. Some animals, as they approached satiation, tasted first one and then the other fluid—a pattern of behavior which was observed also during the first minutes of testing. Such sampling of the fluids appears to be of an exploratory nature and not a true indication of preference.

TABLE IV

PREFERENTIAL TRENDS WITH PAIRS OF SUCROSE SOLUTIONS AND REVERSALS OF PREFERENCES WITH THE ADDITION OF QUININE TO THE PREFERRED SOLUTIONS
($N=24$ rats)

No. of rats showing preferences

Solutions	Tests without *Q*				Concentration of *Q* added to higher concentration of *CHO*								
					7	8	9	10	11	12	13	14	15
4 *CHO*	22	23	24	24	22	23	24	22	20	11	6	2	2
1 *CHO*	2	1	0	0	2	1	0	2	4	13	18	22	22
8 *CHO*	18	23	23	24		24	24	23	19	8	4	1	2
1 *CHO*	6	1	1	0		0	0	1	5	16	20	23	22
16 *CHO*	23	24	24	24		24	24	24	24	20	12	3	2
1 *CHO*	1	0	0	0		0	0	0	0	4	12	21	22
32 *CHO*	24	24	24	24		24	24	24	24	23	16	3	1
1 *CHO*	0	0	0	0		0	0	0	0	1	8	21	23
8 *CHO*	19	23	24	24		18	18	11	6	2	1	0	0
4 *CHO*	5	1	0	0		6	6	13	18	22	23	24	24
16 *CHO*	20	21	24	24		22	21	19	11	3	0	0	0
4 *CHO*	4	3	0	0		2	3	5	13	21	24	24	24
32 *CHO*	22	20	23	24		21	21	16	14	3	1	0	0
4 *CHO*	2	4	1	0		3	3	8	10	21	23	24	24

In at least one-fourth of the available records, however, rats turned to 1 *CHO* *after* they had ingested considerable quantities of 32 *CHO* and continued drinking 1 *CHO* until they became inactive. This pattern of behavior suggests that, as rats approach satiation on 32 *CHO*, some of them express a transient preference for a fluid containing less *CHO* and more water. The transient preference is a half-way station on the road to satiation.

(4) *Reversal of preference.* The results summarized in Tables III and

IV show that a preference for the higher concentration of sucrose can be reversed by adding quinine to the preferred fluid. From the data it is possible to determine how much *Q* must be added to a preferred *CHO*-solution to make the compound isohedonic to a simple *CHO*-solution of lower concentration. For a group of rats, the isohedonic concentration is defined as that concentration of *Q* at which half of the animals prefer *CHO* + *Q* and half prefer a simple *CHO*-solution of lower concentration.

The exact location of isohedonic concentrations was determined graphically, by linear interpolation, from the data in Tables III and IV. Similarly, isohedonic concentrations were determined from the data in Table I. Results were read from right to left until the highest isohedonic values were reached. In the test with 1 *CHO* vs. 1 *CHO* + *Q*, the isohedonic concentration was below Q_6 and, unfortunately, could not be determined from the data at hand.

TABLE V

Isohedonic Concenrtations Based upon Group- and Individual Data

Preference test	Group	Individual rats		
		Range	*M*	*N*
1 *CHO* vs. 1 *CHO*+*Q*	<0.00049	0.00024–0.047	0.0037	24
1 *CHO* vs. 2 *CHO*+*Q*	0.047	0.0029 –0.188	0.047	12
1 *CHO* vs. 4 *CHO*+*Q*	0.029	0.0078 –0.094	0.031	23
1 *CHO* vs. 8 *CHO*+*Q*	0.026	0.012 –0.094	0.033	24
1 *CHO* vs. 16 *CHO*+*Q*	0.062	0.023 –0.13	0.059	22
1 *CHO* vs. 32 *CHO*+*Q*	0.081	0.031 –0.25	0.08	23
4 *CHO* vs. 4 *CHO*+*Q*	0.00009	0.00009–0.0059	0.00098	12
4 *CHO* vs. 8 *CHO*+*Q*	0.007	0.0015 –0.063	0.0078	24
4 *CHO* vs. 16 *CHO*+*Q*	0.015	0.0029 –0.047	0.0140	24
4 *CHO* vs. 32 *CHO*+*Q*	0.018	0.0015 –0.094	0.018	24

The isohedonic concentrations for the group are shown in Table V. The values shown are for two sets of tests—with 1 *CHO* as a standard, and with 4 *CHO* as a standard.

Table V also gives, for individual rats, the range and mean isohedonic concentrations, based upon tongue-contacts. For an individual rat, the isohedonic concentration is defined as that concentration of *Q* in the compound solution which yields the same number of tongue-contacts as a standard *CHO* solution of lower concentration. The isohedonic concentrations usually fell between two *Q*-values and were then arbitrarily placed at the midpoint.

In some tests with 1 *CHO* as a standard, it was impossible to determine isohedonic concentrations for all rats because the presence of *Q* inhibited the drinking of both fluids to such an extent that insignificant numbers of licks were obtained. For this reason, we have shown at the right of Table V the numbers of rats upon which the individual values are based.

Individual isohedonic concentrations have been plotted in Fig. 1. There are two sets of values—with 1 *CHO* as a standard and with 4 *CHO* as a standard.

Guilford proposed that the term *isohedon* be employed to designate a contour-line of equal affective value on a stimulus-surface, such as a triangular surface with white, black, and yellow at the corners.[6] The term was coined by analogy to contour-lines for equal barometric pressure (isobars) on weather maps, and the like. The two curves shown in Fig. 1 are isohedons in a stimulus-area defined by the concentration of sucrose and the concentration of quinine hydrochloride. Presumably, all compound solutions represented by an isohedon are hedonically equal to each other and to the *CHO*-standard.

Although some of the points on the 1-*CHO* isohedon do not fit the curve very well, there can be no doubt about the objective existence of isohedonic concentrations. At all levels of *CHO*-concentration, points on the 1-*CHO* isohedon are higher on the *Q*-scale than corresponding points on the 4-*CHO* isohedon. This obviously means that more *Q* must be added to a compound solution to bring the hedonic intensity down to the level of 1 *CHO* than the level of 4 *CHO;* or, to state the matter differently, for a given concentration of *Q* more *CHO* must be present in the compound solution to raise the hedonic intensity to the level of 4 *CHO* than to the level of 1 *CHO*.

DISCUSSION

Negative hedonic value of quinine. In the above studies with quinine hydrochloride and distilled water, we found no evidence that quinine, at any concentration, is preferred to water. The hedonic value of quinine is consistently negative. In this respect quinine hydrochloride differs from NaCl, saccharin, and other substances which are hedonically positive at low concentrations and become negative at higher concentrations. Our result agrees with the findings of Benjamin.[7] It fails to support the hypothesis of Schneirla that all sensory excitations lead to approach-reactions at low intensities and to withdrawal at high intensities.[8]

---

[6] J. P. Guilford, A study in psychodynamics, *Psychometrika*, 4, 1939, 1-23; System in the relationships of affective value to frequency and intensity of auditory stimuli, this JOURNAL, 67, 1954, 691-695.

[7] R. M. Benjamin, Cortical taste mechanisms studied by two different test procedures, *J. comp. physiol. Psychol.*, 48, 1955, 119-122.

[8] T. C. Schneirla, An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal, in the *Nebraska Symposium on Motivation*, 1959, 1-42.

In another study, Kappauf, Burright, and DeMarco found that the palatability level of a sucrose solution (of relatively high concentration) can be raised by addition of a small quantity of quinine.[9] This enhancement-effect must be attributed to some form of interaction between quinine and sucrose, since quinine by itself has no positive hedonic value. In any event, compound solutions that contain both sucrose and quinine may be either positive or negative depending on relative concentrations. The use of compound solutions makes it possible to study the interaction of positive and negative components.

Relativity of aversion-thresholds. Benjamin determined aversion-thresholds for quinine hydrochloride with two procedures: (1) A 24-hr. intake-test with tap water in one bottle and a solution of *Q* in the other. (2) A 1-hr. drinking-test with a single bottle of *Q*-solution presented immediately after a fluid-deprivation of 16 hr.[10] With both procedures, Benjamin presented *descending* series of concentrations whereas in the present work only *ascending* series were used. His aversion-thresholds, expressed here in gm./100 ml. solution, were in the range from 0.0001 to 0.007. The mean for the 24-hr. test was 0.0012, and, for the 1-bottle drinking test, 0.0016. To facilitate comparison with our results, Benjamin's values are plotted in Fig. 1. His range of aversion-thresholds overlaps ours, but in general his values are lower.

Koh and Teitelbaum determined 'absolute' behavioral taste-thresholds in the rat.[11] They employed a high degree of extrinsic motivation. Four rats were trained to discriminate between quinine solutions and distilled water to avoid a shock, and four hungry rats were trained to discriminate between taste-solutions to obtain food. The combined range of quinine-thresholds for these eight rats, expressed in gm./100 ml. solution, was from 0.0003-0.0006, and the combined mean was 0.0005. These values, as shown in Fig. 1, are definitely lower than those obtained in the present study and by Benjamin.

There can be no doubt that aversion-thresholds vary widely from experiment to experiment and with experimental conditions. Various factors might explain the above differences: whether the quinine is presented in solution with distilled water or tap water or in a sucrose-solution; whether there is extrinsic motivation (like that provided by hunger or shock), or whether the animals are need-free and nondeprived; whether the animals are tested continuously for 24 hr. or given a test for 1 hr. or less; whether ascending or descending series are used; and so forth. With diverse procedures, it is not surprising that diverse results are obtained.

---

[9] W. E. Kappauf, R. G. Burright, and W. DeMarco, Sucrose-quinine mixtures which are isohedonic for the rat, *J. comp. physiol. Psychol.*, 56, 1963, 138-143.

[10] Benjamin, *op. cit.*, 119-122.

[11] S. D. Koh and P. Teitelbaum, Absolute behaviorial taste thresholds in the rat, *J. comp. physiol. Psychol.*, 54, 1961, 223-229.

In a previous review of studies upon preferential taste-thresholds for sodium chloride, the senior author pointed out that preferential thresholds are a function of *motivating* conditions rather than an index of receptor-sensitivity.[12] A similar conclusion is justified here. The aversion-thresholds to quinine hydrochloride are clearly relative to experimental conditions.

Preferential trends. In the preliminary tests with simple sucrose solutions, we found that the preference for a higher *CHO*-concentration was immediately apparent when the *difference* in concentrations was large, but that the rats had to learn to make a preferential discrimination when the difference was small. This results recalls an experiment by Young and Asdourian in which a 1% NaCl solution was used as a standard in preference-tests with comparison-solutions of 2, 6, 18, 54% sucrose.[13] The group as a whole, as well as individual rats, showed marked preferential trends with repeated tests. When the *CHO*-concentrations were low, the trends were gradual; when the *CHO*-concentrations were high, the trends were steeper. It was concluded that the rate of growth of a preferential discrimination is a function of the *difference* in palatability between the test-fluids, and that this rate of growth depends upon two factors: (1) the relative intensities of hedonic processes evoked by the test-fluids; and (2) the number of repetitions of a choice. In the present study, as well as in that by Young and Asdourian, the implications for the theory of reinforcement are clear. *Hedonic* differences are the main factors in reinforcement.

Isohedonic concentrations. Perhaps the most important result of the present study is the demonstration that isohedons can be plotted with need-free rats in the sucrose-quinine stimulus area. By varying the concentrations of the two components, one can obtain groups of stimuli that are hedonically equivalent, but that differ in sensory quality. This is one of the main values in the use of compound taste-solutions. Earlier work with standard foods (solids and liquids) revealed transitive series, preferential trends, mutations of preference, and the like, but current work with compound solutions reveals isohedonic contours. This is a big step forward in the quantitative and objective study of affective processes.

---

[12] Young, The role of hedonic processes in motivation, in *Nebraska Symposium on Motivation*, 1955, 193-238.

[13] P. T. Young and D. Asdourian, Relative acceptability of sodium chloride and sucrose solutions, *J. comp. physiol. Psychol.*, 50, 1957, 499-503.

CONCLUSIONS

(1) In a series of preferencial tests with distilled water and solutions of quinine hydrochloride, no concentration of quinine was found to be consistently preferred to distilled water by need-free (nondeprived) rats. As far as any preference developed, it was a preference for distilled water based upon an aversion to quinine rather than upon any positive incentive value. The higher concentrations of quinine tended to inhibit the intake of both test-fluids.

(2) Aversion-thresholds were determined by adding quinine hydrochloride, step by step, to one member of a pair of sucrose-solutions. The mean thresholds varied with level of sucrose concentration in the range from 0.001 to 0.006 gm./100 ml. solution. The thresholds of aversion to quinine are a function of experimental conditions and not an index of receptor-sensitivity.

(3) In a series of preference tests with simple sucrose solutions, preferential trends were obvious when the difference between sucrose-concentrations was small. The preference for the higher concentration appeared immediately when the difference between concentrations was large.

(4) A preference for the sucrose-solution with higher concentration can be reversed by adding quinine, step by step, to the preferred fluid. The data make it possible to determine what concentration of quinine must be added to make a preferred solution isohedonic to a standard solution of lower concentration. With sucrose standards of 1 and 4 gm./100 ml. solution, we obtained data for plotting isohedonic contours in a stimulus-area defined by the concentrations of sucrose and quinine.

PRODUCTION OF CONSTRAINED ASSOCIATES AND THE INFORMATIONAL UNCERTAINTY OF THE CONSTRAINT

By ROGER N. SHEPARD, Bell Telephone Laboratories, Incorporated

When a subject attempts to list, as rapidly as possible, words falling in some specified category (*e.g.* well-known artists), his rate of listing words begins at a relatively high and stable level and then gradually declines as he progressively exhausts the specified category. Indeed, Bousfield and Sedgewick concluded that the rate of producing words at any given time is proportional to the number of words in the specified category that have not yet been produced.[1] At least on the average, this relation may hold across different categories as well as within the same category. Certainly the initial rate of producing words from a large category (like 'famous people') is generally greater than the initial rate for a smaller category (like 'famous artists').

Now, the uncertainty associated with N equiprobable alternatives has been defined as $\log_2 N$. Usually, however, the category is so large that a determination of the number of words, N, by complete exhaustion of the category is not feasible. To the extent, however, that the initial rate of listing is determined by N, that number, and hence the uncertainty, could be estimated from the initial rate. For example, the degree of statistical constraint exerted upon a given word by its surrounding context could be determined simply by (a) deleting the given word from its context and (b) measuring the initial rate at which *S*s produce guesses as to what the missing word might be. Unfortunately, however, the alternative words are not usually equiprobable; some artists are more famous than others, for example, and would therefore tend to be listed first. Accordingly, this study attempts a combined theoretical and experimental examination of how initial rate depends upon uncertainty when the variation that exists in the probabilities of the alternative words is taken into account.

METHOD

To examine experimentally the relation between uncertainty and rate, *S*s were required to list words to go in a blank representing a word deleted from text. The

---

* Received for publication September 16, 1960. The experimental portion of this study was carried out at Harvard University. The theoretical analysis and preparation of the present report were carried out at the Bell Telephone Laboratories. The author is indebted to George A. Miller for a number of important suggestions.

[1] W. A. Bousfield and C. H. Sedgewick, An analysis of sequences of restricted associative responses, *J. gen. Psychol.,* 30, 1944, 149-165.

uncertainty was varied by including different amounts of the context that had originally surrounded the deleted word. The presumption was that, as the amount of context was increased, the number of alternative words that would fit meaningfully in the blank would decrease. The uncertainty for each amount of context was independently estimated, however, by means of the letter-guessing technique introduced by Shannon.[2]

Subjects. The *S*s were 18 students at Harvard University.

Materials. Thirty-six different newspaper passages containing a minimum of proper names and numbers were collected from back issues of the New York *Times*. A table of random numbers was then used to select one word from each of these passages. Next, for each of these 36 randomly chosen words, six cards were prepared giving a blank (representing the selected word) and 1, 2, 4, 6, 10 or 40 words of context just as they occurred (with punctuation) next to the selected word in the original newspaper passage. In the case of the single word of context, that word was selected equally often from the left and from the right of the word represented by the blank. In all other cases an equal number of words of context were taken from the left and right of the missing word. A counterbalanced design insured that each of the 36 words with its surrounding passage was used in the construction of each of the six contextual conditions just once. Finally, the total number of cards was brought to 234 by adding to these 216 cards 18 additional cards each containing just a blank with no words of context.

Word-listing procedure. Each *S* was presented in random order 13 of the prepared cards: one for the O-context and 2 for each of the other contextual conditions. No two cards presented to any one *S* were constructed from the same passage. Furthermore, each card was used just once in the course of running all 18 *S*s. The source of the original passages and the method of construction of the cards were explained to the *S*s. They were then given 5 min. for each card in which to list all the words of which they could think that might have been the missing word represented by the blank. In O-context, this amounted simply to listing all words that might occur in a newspaper. The number of words listed for each card during each ½-min. subinterval of the 5-min. period was recorded.

Letter-guessing procedure. To obtain a measure of uncertainty with which the rate of listing of words could be compared, a letter-guesssing task devised by Shannon was also administered to the 18 *S*s. For half of the *S*s the word-listing procedure preceded the letter-guessing procedure and for half this sequence was reversed. In the letter-guessing procedure, each *S* was presented with 14 of the prepared cards: 2 for each of the 7 contextual conditions. For each card, the *S*s were required to reconstruct the word actually replaced by the blank. This was done by having them make successive guesses for each letter of the missing word, always moving on to the next letter after making a correct guess. They were told that at the end of the word the correct guess would be "space." The restrictions governing the assignment of cards to *S*s in the word-listing procedure also prevailed in the letter-guessing procedure. In addition, for any one *S*, no card used in one of the procedures was constructed from the same passage as any card in the other procedure.

---

[2] C. E. Shannon, Prediction and entropy of printed English, *Bell Syst. tech. J.*, 30, 1951, 50-64.

Calculation of the uncertainty. The uncertainty was estimated for each of the seven contextual conditions on the basis of the distribution of the number of guesses required on the average to reconstruct the missing word. The method of calculation was basically that used by Shannon to establish an upper bound on the entropy of printed English.[8] Since the missing words varied in length, however, the uncertainty for letters beyond the position containing the space following the shortest word (viz. the third position) were calculated by considering that after correctly guessing "space" any

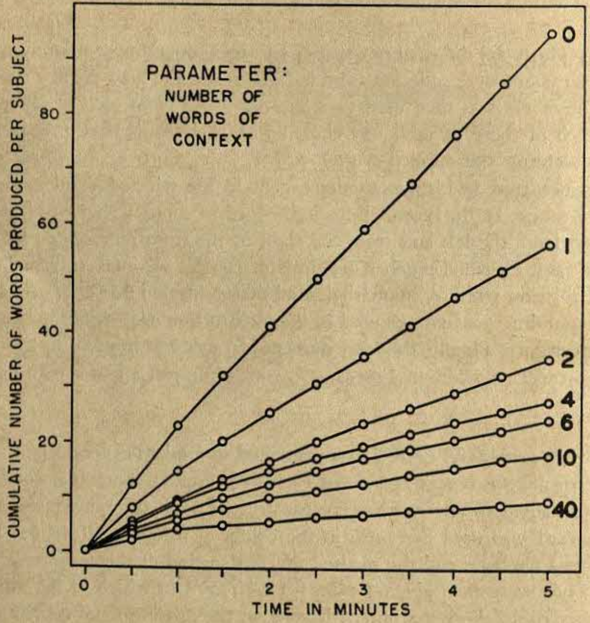


FIG. 1. CUMULATIVE NUMBER OF WORDS PRODUCED PER *S* UP TO EACH HALF-MINUTE SUBDIVISION OF THE FIVE-MINUTE PERIOD

A separate empirical curve is plotted for each of the seven amounts of context supplied.

---

[8] Shannon, *op. cit.* The actual method of calculation was slightly modified in the present application. First, the smoothing operation described by Shannon on p. 64 was not used. Secondly, since the uncertainty decreases rapidly only for the first few letters of a word, it was calculated separately for the letters in the first, second, and third positions in the missing word; but the distributions of guesses were combined before calculating the uncertainties for letters in the fourth and fifth positions and for letters in the sixth through tenth positions.

further guesses would also be correct (for the *S* could then have no residual uncertainty about the missing word). Adherence to this rule insures that the calculated uncertainty is always zero for any position beyond that containing the space following the longest word. The total uncertainty for each context condition was obtained by adding together the uncertainties estimated for letters in the first, second, etc., positions in the missing word.

RESULTS

In Fig. 1 the average number of words listed per *S* is cumulatively plotted as a function of time for each of the seven contextual conditions. The results are consistent with the notion that the inclusion of more context further limits the number of available alternatives from which the missing word might be drawn and therefore decreases the rate of listing those alternatives.

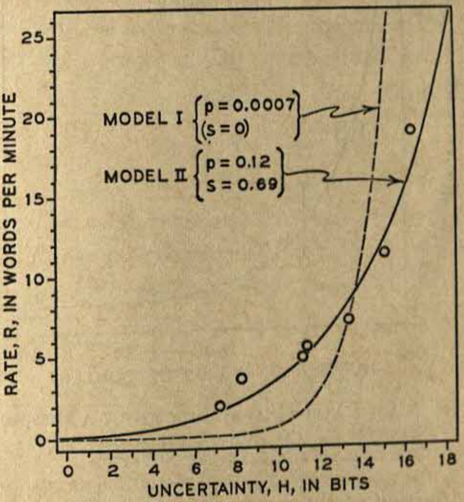


Fig. 2. Average Rate of Producing Words during the Five-Minute Period as a Function of the Estimated Uncertainty about the Missing Word

The highest point is for words of 0-context and the lowest point is for 40 words of context. The broken curve is based on the assumption that all of the alternative words are equiprobable, and the solid curve on the assumption that the probabilities conform to a Zipf-like distribution.

In Fig. 2, the average rate of producing words during the first 5 min. is plotted for each contextual condition against the uncertainty as estimated from the letter-guessing procedure. The two fitted curves will be considered in the following theoretical discussion. It is sufficient to note here that the solid theoretical curve (labeled "Model II") provides a reasonably good approximation to the empirically obtained relation between rate and uncertainty.

To the extent, then, that the solid theoretical curve is accepted as a satisfactory approximation, that curve can be used to convert an observed rate into an estimate of uncertainty. An application of this possibility is illustrated in Fig. 3, where estimated uncertainty about the missing word

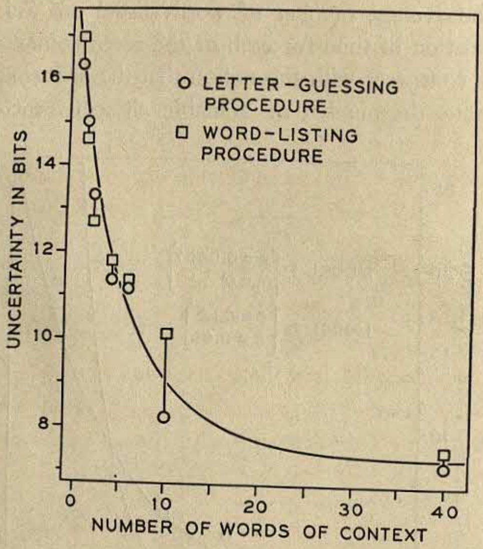


FIG. 3. ESTIMATED UNCERTAINTY ABOUT A MISSING WORD AS A FUNCTION OF THE AMOUNT OF BILATERAL CONTEXT SUPPLIED

is plotted as a function of the number of words of context supplied. The values represented by the circles are the estimates calculated directly from the data of the letter-guessing procedure. The values represented by the squares, however, were obtained by converting the rates of listing words into uncertainties via the solid theoretical curve. With the possible exception of the 10-word context, the two procedures yield rather similar results. For convenience of interpolation, therefore, an exponential decay-function has been fitted to the combined data.

THEORETICAL DISCUSSION

In view of the rather orderly relations found empirically, between the independent and dependent variables of Figs. 1 and 2, the question might usefully be raised as to whether there exist some simple assumptions that would predict essentially these obtained relations. As a first step in this direction, the following three simplifying assumptions seem reasonable candidates. The first is that if a word has not yet been produced, the probability that it will be produced during the next unit of time does not depend upon how long the experiment has been in progress or upon what other words have already been produced. The second is that this probability drops to zero for any word as soon as that word is produced (*i.e.* no word is produced twice). The third is that the time required to produce a word is negligible. These assumptions are undoubtedly violated in actual cases, but such violations hopefully represent second-order effects and do not alter the general features of the relations considered here.

Suppose, then, that p_r denotes the probability that the r^{th} of the N alternative words will be produced during the next unit of time if that word has not yet been produced. The initial rate of listing words is then simply

$$R = \sum_{r=1}^{N} p_r \tag{1}$$

Moreover, if just one word is to be selected out of the N, the probability that it will be the r^{th} can now be expressed as p_r/R. According to Shannon's definition,[4] the uncertainty associated with the set of N alternative words is therefore given by

$$H = - \sum_{r=1}^{N} (p_r/R) \log_2 (p_r/R). \tag{2}$$

This equation shows that the relation between uncertainty, H, and rate, R, depends upon the distribution of the probabilities, p_r. Accordingly, two kinds of distributions will now be considered.

Model I: Rectangular distribution of the probabilities: If all N words are equiprobable, $p_r = p$ for $r \leq N$, zero otherwise. Hence the predicted relation between uncertainty and rate is simply

$$H = \log_2 (R/p). \tag{3}$$

This equation is represented by the broken curve in Fig. 2. Clearly, the assumption of equiprobable words is not adequate even as a rough approximation. Moreover, to achieve even this degree of fit, the probability of producing any one word during a 1-min. interval must be assigned an implausibly small value (viz. $p =$ 0.0007).

---

[4] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, 1949.

Model II: Zipf-like distribution of the probabilities: Evidence assembled by Zipf and others suggests that the distribution of the probabilities, far from being rectangular, is generally approximated more closely by setting

$$p_r = pr^{-s} \quad \text{for} \quad r \leq N, \tag{4}$$

and zero for $r > N$. Here, p is the probability per unit time for the production of the most probable word ($p = p_1$) and, since the words are considered to be ranked according to probability (so that $p_1 \geq p_2 \cdots \geq p_N$), s is a slope constant governing how steeply the probabilities for the other words drop away from this maximal value.[5]

Except when N is very small, the sums in Equations [1] and [2] will be closely approximated in this model by the corresponding definite integrals. Integration and simplification then yields

$$R \cong p(N^{1-s} - 1)/(1 - s) \tag{5}$$

and

$$H \cong \log_2 \frac{R}{p} + \left[s + \frac{sp}{(1-s)R} \right] \log_2 N - \frac{s}{(1-s)} \log_2 e. \tag{6}$$

If Equation [5] is used to eliminate N from Equation [6], the final result can be put in the form

$$H \cong \log_2 \frac{R}{p} + \frac{s}{1-s} \left\{ \frac{p}{R} \left[\frac{R}{p} + \frac{1}{1-s} \right] \log_2 \left[(1-s) \frac{R}{p} + 1 \right] - \log_2 e \right\}$$

Since Model II reduces to Model I, as $s \rightarrow O$, the rectangular distribution of the probabilities can be viewed as a special case of the Zipf-like distribution. But, whereas in Model I each distribution is governed by just two parameters (p and N), in Model II each distribution is governed by three parameters (p, s, and N). For simplicity, the parameters p and s are assumed to be essentially constant from one situation or context to another. As in Model I, then, variations in rate, R, and uncertainty, H, will be attributed only to variations in the number of alternative words, N. Model II, however, should provide for a closer account of the data since it contains an additional parameter to accommodate differences in the probabilities of the alternative words. That the fit is indeed better is shown by the solid curve in Fig. 2. Moreover the estimated probability (per minute) of the most likely word now rises to a more plausible value (viz. $p = 0.12$).

Model II also provides a better account for certain other phenomena. For example, Model I can easily be shown to imply that the cumulative curves plotted in Fig. 1 should conform to the negatively accelerated exponential functions originally proposed by Bousfield and Sedgewick.[6] But these curves (as well as some presented by Bousfield and others) actually rise more rapidly at first and then more slowly than the best-fitting exponential. This is just the kind of departure that is predicted as soon as the probabilities are asssumed to vary for different words (as in Model II).[7] Also, according to Model II but not Model I, some words will have higher probabilities of production per unit time and hence on the average will (a) be produced by more *S*s during a fixed period of time and (b) be produced earlier by any one *S*. Such a relation between frequency and latency of associative responses is already known under the name of Marbe's law and appears also in the negative correlation (found by Bousfield and Barclay) between the frequency of listing a word and its mean rank order of appearance in the list.[8]

---

[5] G. K. Zipf, *Human Behavior and the Principle of Least Effort*, 1949; 1-573; Murray Aborn and Herbert Rubenstein, Perception of contextually dependent word-probabilities, this JOURNAL, 71, 1958, 420-422; Benott Mandelbrot, On recurrent noise limiting coding, *Proceedings of the Symposium on Information Networks, Polytechnic Institute of Brooklyn*, 1954, 205-221; G. A. Miller, Some effects of intermittent silence, this JOURNAL, 70, 1957, 311-314; *Language and Communication*, 1951, 90-93; H. A. Simon, On a class of skew distribution functions, *Biometrika*, 42, 1955, 425-440; B. F. Skinner, The distribution of associated words, *Psychol. Rec.*, 1, 1937, 69-76.

Although Model II is an improvement over Model I, it is itself put forward only as a tentative approximation to the highly complex processes studied here. In fitting the model, for example, no account was taken of the fact that different contexts do not exert the same degree of constraint simply because they contain the same number of words. Also, when there is little or no context, the simplifying assumption that the probability of producing a given word is independent of what words have already been produced is certainly inaccurate. When a given word (*e.g.* 'table') is produced, the probabilities of other associated words ('chair,' 'desk,' etc.) are undoubtedly increased temporarily over their usual level.[9] Such associative chaining may have spuriously augmented the rate, particularly for the O-context condition.[10]

---

[6] For the kind of derivation required see, for example, R. N. Shepard, Stimulus and response generalization: Deduction of the generalization gradient from a trace model, *Psychol. Rev.,* 65, 1958, 242-256, especially 248.

[7] For a discussion of this point (but in somewhat different connections) see W. K. Estes and C. J. Burke, A theory of stimulus variability in learning, *Psychol. Rev.,* 60, 1953, 281-282; or R. N. Shepard, Applications of a trace model to the retention of information in a recognition task, *Psychometrika,* 26, 1961, 185-204. A stronger test of Model II would be possible if functions could be derived from the model and fitted to the cumulative curves in Fig. 1. Unfortunately, attempts to obtain a closed expression for these curves have not been successful. However, certain polynomial approximations to the exact expressions have been examined. These suggest that (although Model II does come closer than Model I) even Model II cannot account for the full degree of curvature found in the early segments of the cumulative curves.

[8] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology,* 1954, 61-65; W. A. Bousfield and W. D. Barclay, The relationship between order and frequency of occurrence of restricted associative responses, *J. exp. Psychol.,* 40, 1950, 643-647. See, also, the more recent paper: W. A. Bousfield, B. H. Cohen, and J. G. Silva, The extension of Marbe's law to the recall of stimulus-words, this JOURNAL, 69, 1956, 429-433.

[9] W. A. Bousfield, The occurrence of clustering in the recall of randomly arranged associates, *J. gen. Psychol.,* 49, 1953, 229-240; J. J. Jenkins, W. D. Mink, and W. A. Russell, Associative clustering as a function of verbal association strength, *Psychol. Rep.,* 4, 1958, 127-136; J. J. Jenkins and W. A. Russell, Associative clustering during recall, *J. abnorm. soc. Psychol.,* 47, 1952, 818-821; R. N. Shepard, Multidimensional scaling of concepts based upon sequences of restricted associative responses, *Amer. Psychologist,* 12, 1957, 440-441 (abstract).

[10] The method used for estimating uncertainty is not free from bias, either. Strictly, it represents only an upper bound. But, since the data were not smoothed before calculating the uncertainty (see footnote 3 *supra*), the obtained estimates are somewhat below the upper bound. Also, there is reason to believe that, when there is little context to control their associations, *S*s do not behave like Shannon's "ideal predictor" (Fred Attneave, Psychological probability as a function of experienced frequency, *J. exp. Psychol.,* 46, 1953, 81-86).

Despite these difficulties, the results are generally in rough agreement with those obtained previously by very different methods. The value estimated here for the slope parameter governing the distribution of probabilities (viz. about 0.7 in Model II) is smaller than the values near 1.0 typically found by Zipf. Zipf was not concerned, however, with the probabilities that *S*s would produce words to go in given contexts, but with the frequencies of actual occurrences of those words (without regard to context) in printed English texts.[11] Rothkopf, who recently had *S*s produce words subject to certain constraints, obtained estimates of the slope parameter, s, ranging from about 1.0 down to 0.76. His constraints were not, however, contextual in the present sense.[12] The average size of the total vocabulary can be estimated for the present *S*s by substituting for the variables p, s, and R in Equation [5]. The estimate for N obtained in this way is extremely unreliable; the 20% deviation of the empirically determined rate for the O-context from the solid theoretical curve results in a 150% deviation of the estimate for the size, N, of the vocabulary. If the point of the solid theoretical curve corresponding to O-context is taken as the more stable estimate of R, N is estimated as 174,000 words. This seems large for the size of a writing vocabulary. Since the *S*s were not asked, however, to justify or define any of the words they produced, the number of words that can be recognized (rather than the number that are actually used in writing) may be the more appropriate figure for comparison. The present estimate for N does fall within the range of about 100,000 to 200,000 that has been estimated for the size of the recognitive vocabulary.[13] However, these conclusions are extremely tentative in view of the instability of the estimate for N and the factors (such as associative chaining) that undoubtedly perturb any estimates based upon O-context.

The uncertainty about a single word when no context is supplied was estimated here as 16.3. Since there was an average of 6.3 letters per word (including the space), this yields an estimate of 2.59 bits per letter. This is to be compared with Shannon's estimate of 2.62 bits per letter for words selected at random from printed English.[14] At the other extreme, the average uncertainty estimated here for the 40-word context can be expressed as 7.20/6.3 or 1.14 bits per letter. This is smaller than the upper bound of about 1.3 bits per letter (with 100 letters of context) estimated by Shannon, or that of about 1.9 bits per letter (with 10,000 letters of context) estimated by Burton and Licklider.[15] The contexts in those studies were, however, supplied only on one side of the missing material. Aborn, Rubenstein, and Sterling; Miller and

---

[11] Zipf, *op. cit.*, 25.

[12] E. Z. Rothkopf, Two predictors of stimulus equivalence in paired-associate learning, *Psychol. Rep.*, 7, 1960, 246. It should also be noted here that, whereas Model II applies strictly only when R is interpreted as the initial rate, the rates actually used in estimating s were the average rates for the first 5 min. Since the curves in Fig. 1 are slightly convex, the initial rates must be higher than those estimated. Although the data become less reliable, a reasonably good fit of Equation [7] can still be achieved when the data for only the first minute are used. The most satisfactory fit then results when p is increased from 0.12 to about 0.27 and when s is increased from 0.69 to about 0.77.

[13] For a discussion of studies of vocabulary-size see, for example, G. A. Miller, *Language and Communication*, 1951, 121 and 149.

[14] Shannon, *Bell Syst. tech. J.*, *op. cit.*, 53.

[15] N. G. Burton and J. C. R. Licklider, Long-range constraints in the statistical structure of printed English, this JOURNAL, 68, 1955, 650-653.

Friedman; and Garner and Carson have all shown that a bilaterally distributed context (as supplied in the present study) exerts greater constraint than a unilateral one.[16] Except for this difference, the curve plotted in Fig. 3 is analogous to those presented by Shannon and by Burton and Licklider. This curve agrees with their data as well as those of Aborn, Rubenstein, and Sterling and of Garner and Carson in suggesting that only negligible constraint is exerted over spans exceeding 10 words (or about 50 consecutive letters) of text. (This corresponds to the 20-word point on the abscissa of Fig. 3, since the context was supplied bilaterally.)

The method proposed here for estimating the uncertainty about a missing word from the rate of listing alternatives differs in three respects from the previously used letter-guessing method devised by Shannon. First, words rather than letters serve as units. Secondly, it is not necessary to know what the deleted word actually was. And thirdly, the method is not restricted to estimating the average uncertainty for a large sample of different contexts, but can in principle be used for one particular context.[17] On the other hand, the present method has the very serious drawback that the parameters p and s are assumed to be invariant from one context to another. If, for a particular context, p is unusually high, for example, then the uncertainty (as estimated from the rate) would be spuriously inflated.

SUMMARY

*S*s were shown a blank (representing a word selected at random from English text) and either 0, 1, 2, 4, 6, 10, or 40 surrounding words of context. They were asked to list in a 5-min. period as many words as they could that might have been the word replaced by the blank. The major results were as follows:

(1) The average rate of listing words monotonically dropped by a factor of 10 from about 20 words per minute, when no context was provided, to about 2 words per minute, when the maximum amount of context was supplied.

(2) Correspondingly, this rate increased monotonically with the informational uncertainty about the missing word as independently estimated by Shannon's letter-guessing method.

(3) A theoretical account of the relation between uncertainty and rate was proposed by considering that a word that has not been produced up to

---

[16] Murray Aborn, Herbert Rubenstein, and T. D. Sterling, Sources of contextual constraint upon words in sentences, *J. exp. Psychol.*, 57, 1959, 171-180; W. R. Garner and D. H. Carson, A multivariate solution of the redundancy of printed English. *Psychol. Rep.*, 6, 1960, 123-141; G. A. Miller and E. A. Friedman, The reconstruction of mutilated English texts, *Information and Control*, 1, 1957, 38-55.

[17] Shannon's calculation assumes that the distribution of guesses for the n^{th} letter is the one that results when the *S*s (predictors) base each guess on a different preceding sequence of n-1 letters. If the same sequence of n-1 letters were given to all predictors and they behaved like Shannon's "ideal predictor," they would all arrive at the correct letter after exactly the same number of guesses and, hence, his calculation would always yield an estimate of zero bits for the uncertainty.

any given time has a fixed probability of production during the next small interval of time. The agreement between the theoretical account and the empirical data was unsatisfactory when the alternative words were assumed to be equiprobable, but became reasonably satisfactory when the words were assumed to vary in probability in accordance with a Zipf-like distribution.

(4) Since, on the average, the rate of producing words in a specified category is apparently closely related to the uncertainty associated with that category, uncertainty might under some conditions be conveniently inferred from rate. When this procedure was used, for example, to determine how uncertainty about a missing word depends upon the amount of context supplied, the results were consistent with those obtained earlier using other methods.

FACTORS AFFECTING STRATEGIES USED IN PROBLEMS OF CONCEPT-FORMATION

By LYLE E. BOURNE, JR., University of Utah

Several writers have recently discussed the need for and the value of more incisive analyses of human learning than those based on an inspection of such over-all measures of performance as trials, errors, or time to criterion. Such indices provide a precise measure of general efficiency of performance but indicate little of the plan, strategy, or heuristic employed by *S* while mapping a sequence of responses.[1] There is presently a lack of evidence on the nature of strategies used by *S*s in learning situations and on their relationship to important variables of the tasks and, thus, little upon which an empirical evaluation of their informativeness as measures of performance can be based. The purpose of this experiment was to explore the effects of four major variables, *viz.* amount of stimulus-variation, type of stimulus-variation, method of stimulus-presentation, and complexity of task upon both the over-all efficiency and the strategies of *S*s in concept-problems.

The method of this experiment was patterned after that of Bruner *et al.* which appears to be well-suited to the study of strategies employed by *S*s in concept-formation.[2] *S*, whose task was to discover a concept about a set of geometric figures, gave his best guess (hypothesis) about the concept after the presentation of each stimulus. *E* determined the order in which the stimuli were presented and provided *S* with knowledge of the correctness of each hypothesis. Two general strategies of the *reception*-type, described in detail by Bruner *et al.*, were available to *S*. Briefly, the *S*s using the *wholist* strategy stated all attributes of the first positive instance of the concept as a working hypothesis and the *partist* *S*s used

---

* Received for publication August 21, 1961. This research was supported in part by a grant (M-1759) from the National Institute of Mental Health, U.S. Public Health Service. Data were collected during the summer, 1960, at the University of Wisconsin. The author wishes to express his gratitude to the staff of the Department of Psychology, University of Wisconsin, for their support and for the use of their facilities.

[1] J. S. Bruner, J. J. Goodnow, and G. A. Austin, *A Study of Thinking*, 1956, 1-80; G. A. Miller, Eugene Galanter, and K. H. Pribram, *Plans and the Structure of Behavior*, 1960, 1-39; Allen Newell, J. C. Shaw, and H. A. Simon, Elements of a theory of human problem solving, *Psychol. Rev.*, 65, 1958, 151-166.

[2] Bruner, Goodnow, and Austin, *op. cit.*, 81-230.

some number of attributes less than the total. Both strategies required appropriate modifications of the initial hypothesis as information about relevant and irrelevant attributes of the stimulus became available with successive trials.

In this experiment, amount of stimulus-variation was quantified in terms of the degree to which each stimulus was different from its immediate predecessor, *i.e.* in terms of the number of attributes of the stimulus which changed from one trial to the next. Although there is a lack of experimental data, it was expected that the difficulty of the task would increase directly with amount of variation, assuming that large amounts of variation interfere with locating and using relevant information. At least two basic types of stimulus-variation are conceivable in this task; both were used. One type assumes the first positive instance as a focus for changes in the attributes of all successive stimuli. Variation of the second type is made with a shifting focus, the attributes of each successive positive instance serving as a basis of subsequent variation. Experimental details will be described later; however, it should be mentioned that the sequences of stimuli generated under these two conditions are similar to those *S*s have been observed to select voluntarily in some earlier studies. Bruner *et al.*, for example, reported that 'ideal' *S*s order selections of stimuli around some fixed focal instance, generally the first positive instance, while less efficient *S*s appear to shift focus periodically.[3] Types of variation studied in the present experiment were not designed, however, as objective manipulations of strategies of selection; the variable was introduced into the design of the experiment primarily to obtain more general evidence on the effects of amount of variation.

Two methods of presentation of stimuli, similar to those used by other investigators, were studied.[4] In the *simultaneous* condition, all stimuli were arrayed for *S*'s inspection during the entire experimental session; in the *successive* condition, stimuli were displayed one at a time. The latter method, because it required *S* to recall stimuli presented previously, was expected to result in inferior performance.

Complexity of the task was quantified in terms of the number of attributes of the stimuli which defined the concept. Several earlier studies have been concerned with similar variables and have typically shown a direct relationship between complexity and difficulty of the task.[5] Evidence from this study will permit conclusions about the interactive effects of complexity with other major variables and of its influence on the use of strategies of *reception*.

---

[3] J. S. Bruner, G. A. Austin, and R. V. Seymour, Free-choice strategies in concept attainment, *Amer. Psychologist*, 8, 1953, 327 (abstract).

[4] Bruner, Goodnow, and Austin, *op. cit.*, 92-96; C. I. Hovland and W. Weiss, Transmission of information concerning concepts through positive and negative instances, *J. exp. Psychol.*, 45, 1953, 175-182.

METHOD

Stimulus-materials. Each of 64 different geometric designs was printed on a separate stiff, white, 3-in. square cardboard.[6] These designs varied in six binary dimensions, *viz.* form, size, color, texture, number, and border; a particular design portrayed one of two attributes or values within each dimension, *e.g. solid* or *spotted* surface (texture) or *continuous* or *broken* outline enscribing the form (border).

Subjects. Participants in the experiment were 48 volunteers from undergraduate classes in psychology. Each was assigned unsystematically to a condition of the experiment, served individually, and completed his performance within 45 min.

Task. At the outset of a session, the nature of the task was fully described to *S*. Examples of conjunctive concepts that might be formed with the stimulus-cards were shown along with examples of concept-problems. Concepts of all levels of complexity were illustrated, but none of these overlapped the solutions to problems used in the experiment proper.

S was then started on the first of three concept-problems; in no case did any two of the three problems have solutions which overlapped in relevant or correct attributes of the stimuli. The three problems given to each *S* were characterized by three different levels of complexity (number of attributes relevant to the concept). Concepts employed in this study, all conjunctive, were defined by either one, two, or three attributes of the total of six, *e.g.* single-attribute concept—*red* figure; double-attribute concept—*red circle;* triple-attribute concept—*red circle* within a *continuous* border. Method of presenting stimuli, either *successive* or *simultaneous*, was the same for all three of *S*'s problems. In the simultaneous condition, all 64 stimulus-cards were presented in an orderly array and *S* was allowed free inspection throughout the problem. The array used was an 8 × 8 matrix with red cards occupying the first four rows and green the remaining four, circles in the first four columns, ellipses in the remaining. Each dimension was displayed in a prearranged pattern with this general ordering, such that any two neighboring cards differed in one and only one attribute.

In all problems, the first card displayed (successive condition) or pointed to (simultaneous condition) was a positive instance of the concept. To be informative, the second card, and, indeed, all other cards, had to differ in at least one dimension from the first, although it is obvious that this procedure does not insure that every stimulus will provide completely non-redundant information. Three levels of variation of stimuli (*i.e.* number of attributes of the total of six which changed between successive stimuli) were employed in the experiment; successive stimuli differed in one, two, or three attributes. If each stimulus was maximally informative, *S* would have sufficient information to solve a problem (*i.e.* name the concept) in differing numbers of trials under the three conditions. To eliminate this confounding of two potentially important variables, all information necessary to solve a problem was spread over a fixed number of stimuli, seven, in every condition of the experiment. This is possible without the necessity of any stimulus being completely uninformative although it produces differing numbers of positive instances in the first seven stimuli among the various conditions of the experiment. Nonetheless, in all cases there was a sufficient number of positive and negative instances to solve the problem in seven trials and the procedure permitted the required controls over any extraneous systematic trial-by-trial effects.

---

[5] Austin, Bruner, and Seymour, Fixed-choice strategies in concept attainment, *Amer. Psychologist*, 8, 1953, 314 (abstract); L. E. Bourne, Jr. and R. C. Haygood, The role of stimulus redundancy in concept identification. *J. exp. Psychol.*, 58, 1959, 232-238.

[6] The stimulus-materials were developed and kindly provided by E. James Archer of the University of Wisconsin.

Two types of variation of stimuli were used in the experiment. In the *fixed-focus* condition, all stimuli following the first contained attributes identical to the first, except for the one, two, or three attributes which were changed on that trial. All dimensions of the stimuli changed in value or attribute at least once in the sequence of trials, but attributes of the first instance were always recovered or reinstated after change. The *shifting-focus* type of variation was based on the attributes of each successive positive instance, not necessarily the first. In this case, each stimulus-card had attributes identical to the last presented positive instance except for the attributes which were changed on that trial. In effect, this means that only relevant attributes were necessarily reinstated and an irrelevant attribute appearing in the initial instance might never again occur in the stimulus-sequence. It can be shown that under conditions of large variation in stimuli, there is little difference in the sequences generated by these two types; however, markedly different orders obtain when only one dimension is varied from stimulus to stimulus.

One remaining variable was included in the experiment to offset the effects of possible communication among the *S*s about the requirements of the task. Two sets of problems, differing in the attributes which defined the concepts, were used. This variable was held orthogonal to other effects that variance in the data attributable to it could be computed. The attributes of the stimuli defining the two sets of problems were: (a) *red* figure, *solid* textured *ellipse,* and *two small* figures within a *broken* border, and (b) *one* figure, *large* figure within a *continuous* border, and *green spotted circle.* Note that each set consists of three problems, each with a different number of unique relevant attributes, and that all attributes of the stimuli are relevant in exactly one problem.

S was instructed to attend to each stimulus displayed or pointed to by *E.* Each stimulus was designated positive or negative with respect to the concept. *S* was to respond to each card with a statement of his hypothesis about the concept illustrated. Instances were displayed or pointed to in a prearranged order, which conformed to one set of experimental conditions, until *S* correctly stated the concept. All the *S*s were required to work the problems without the aid of paper and pencil, and at no time did *E* supply information requested by *S* about instances presented previously. Each of *S*'s hypotheses was recorded by *E*, but never within a given problem did *S* have recourse to this record. The rate of presentation was determined by *S*, who was allowed as much time as needed to respond to any card. The series of three problems given to *S* was completed with no rest between successive problems.

Design. Fundamentally, the design of the experiment was a $2 \times 2 \times 2$ factorial with repeated measures; however, the principle of the Latin square was introduced to counterbalance the factors of complexity and amount of stimulus-variation across the three problems presented to each *S*. The three factorially arranged variables were method of presentation, type of stimulus-variation, and problems. Within each of the eight groups formed by the factorial were six *S*s, each of whom was run through the series of three problems designated by a row of a 3×3 Greco-Latin square. The same square was used twice, to account for the total of 6 *S*s within each group. Latin letters of the square were identified with the three conditions of complexity and Greek letters with the three degrees of stimulus-variation.

RESULTS AND DISCUSSION

Trials to solution. Many *S*s required more than the seven non-redundant instances to solve a problem; on the other hand, a few *S*s correctly named a concept by chance before all required information had been presented. All correct solutions were accepted and the *S*s were scored on each of three problems on the basis of the number of instances presented before solution. In an analysis of variance performed on these data, several effects were shown to be statistically significant. The large variance associated with method of presentation ($F = 14.66$, 1/40 *df*, $p < 0.01$) served to confirm similar findings of earlier experiments.[7] In all experiments, the *S*s have been markedly handicapped by the necessity to remember the potentially relevant information in previously presented stimuli. Those *S*s who performed their tasks with an orderly array of stimuli before them had a clear advantage. This result points to the need for a more detailed and complete investigation of this variable. The conditions explored here are merely end-points on a quantifiable dimension, which might be operationally defined in terms of the number of previously presented stimuli to which *S* has access on any trial. A condition demanding less memory than the successive mode of presentation would permit *S* to see one previously shown stimulus along with the current one. An even less taxing condition would allow *S* access to two previously shown stimuli. The case in which all stimuli are arrayed for *S* is essentially the same as that in which all the stimuli absolutely necessary to convey the complete solution of the problem are available, *i.e.* *S* has access to all previously shown instances. An experiment incorporating a procedure of variation such as this would provide a more detailed understanding of the variable.

Bruner *et al.* found that the *S*s who voluntarily switched focus in the course of a problem performed in a significantly inferior way to those who did not. Their results suggest that to use relevant information optimally, *S* must order it about a fixed focus; such a procedure minimizes interference in memory.[8] In the present experiment, however, shifting focus of stimulus-variation produced performance significantly superior to that under conditions of fixed focus ($F = 4.20$, 1/40 *df*, $p < 0.05$). An inspection of the sequences of stimuli generated by these types indicated that focus shifting provided *S* with a better running record of potentially important attributes than did the condition of fixed focus. Recall that the shifting condition allowed any change in an irrelevant dimension from one stimulus to the next to remain without reinstatement of the original attribute in later cards. On the other hand, any change that affected a relevant dimension was necessarily changed back to the initial attribute in succeeding instances. All changes regardless of the relevance of the related dimension(s) were not incorporated within succeeding stimuli in the condition of fixed focus since the initially presented card remained the focus throughout the problem. This difference between conditions, resulting in differential treatment of relevant and irrelevant information, probably accounts for the superiority of the condition of shifting focus. A significant interaction of type of variation with method of presentation ($F = 4.20$, 1/40 *df*, $p < 0.05$) indicated that the main influence of shifting focus in the present investigation appeared in the successive condition in which *S* must retain information from previous stimuli. When all cards were simultaneously displayed for *S*, no difference between types of variation was found.

---

[7] Bruner, Goodnow, and Austin, *op. cit.*, 92-96; Hovland and Weiss, *op. cit.*, 178-180.

Number of relevant or defining attributes of a concept determined to a significant degree ($F = 10.76$, 2/61 *df*, $p < 0.01$) the difficulty *S* had in identifying that concept; mean trials to solution increased directly with the number of relevant attributes (*i.e.* complexity). On the other hand, amount of stimulus-variation and all interactions with amount of stimulus-variation were statistically insignificant; the expectation that performance would worsen with increased variation of stimuli was not confirmed. It should be recalled that information required to solve any problem was spread over the first seven instances presented. It is probably the case that amount of stimulus-variation alone is an ineffective variable and that its assumed importance derives from its typically high correlation with amount of information per instance.

Strategies of reception. In an attempt to supplement the analysis of trials, several investigations, patterned after those reported by Bruner *et al.*, were made of the strategies employed by the *S*s. Contrary to earlier findings, the *S*s in the present study predominantly used the *partist* strategy. That is, upon being shown the first card, which in all cases was a positive instance of the concept, most of the *S*s chose as an initial hypothesis only a part of the total number of available attributes of the stimulus. Only 9% of all problems presented were initially attacked by the *wholist* method. This is in contrast to 62% reported by Bruner and his coworkers.[9] The discrepancy may result from the fact that the *S*s in the earlier study were given a total of 14 problems during the course of the experiment; those in the present study were given only 3. As *S* becomes more practiced in the performance required, his selection of a mode of approach should come closer to that which is most efficient, and, as the data of this experiment clearly show, the wholist strategy is more efficient than any partist approach. It is proposed, then, that Bruner's *S*s, because they were more sophisticated, were more likely to display a greater preference for the wholist strategy than those in the present study. It is possible that the nature of sample-problems used in the instructions set *S* to look for simple concepts and that this might account for the high frequency of partists. The results of Bruner *et al.* might then be attributed to the greater opportunity for their *S*s to overcome this set with experience.[10] These explanations, however, are in need of further investigation before either can be fully accepted.

---

[8] Bruner, Austin, and Seymour, *op. cit.*, 327.

It should be noted that the *S*s were categorized as wholists or as partists solely on the basis of the number of attributes included within the *first* hypothesis. Some *S*s expanded the number of attributes under consideration (*i.e.* became more like wholists) later in the sequence of stimuli. It is conceivable also that *S* might keep track of all attributes of the stimuli without reflecting this in his stated hypothesis. Thus the reported number of wholists may be an underestimation of their actual frequency. If the definition of a wholist is liberalized to include *S*s who use all attributes of the stimuli in any hypothesis during the first *three* trials of a problem, their frequency increases to 17%. This procedure results in no essential change, however, in the trends reflected in the following analyses which, in all cases, utilize the more rigorous original definition of a wholist.

Of problems attacked by the wholist method, all but one were solved within seven instances. The mean number of instances required to reach solution for these problems was 7.31, which stands in marked contrast to the mean number of instances required by the partist's strategy, 9.75 ($F = 4.01$, 1/142 *df*, $p < 0.05$). Wholists were also less variable in the number of trials to solution; the standard deviation of trials required by the wholist was 1.38, whereas that of the partist was 4.27. The partists were inclined to guess early and frequently in the course of working on a problem. More often than not these early guesses were incorrect, misleading, and productive of an inordinately large number of trials to solution. Occasionally, however, a guess was correct and served to reinforce strongly *S*'s tendency to repeat such behavior on succeeding problems. Wholists, on the other hand, had a more clear-cut, formalized plan for solution. Deviations from it were rare, guesses infrequent. The plan in all cases required seven instances.

---

[9] Bruner, Goodnow, and Austin, *op. cit.*, 126-155.
[10] Bruner, Goodnow, and Austin, *op. cit.*, 147.

Important differences among *S*s are not clarified by the partist-wholist analysis. There is a continuum, quantifiable in terms of the number of attributes which *S* includes as a part of his initial hypothesis, underlying the classification. The *S*s in this experiment represented every possible size of initial hypothesis (1 through 6 attributes). The question of differences between partists and wholists is basically a question of the relationship between performance and size of initially used hypotheses. As can be seen in Table I, there were consistent changes in trials to solution and percentage of solutions within the first seven trials as size of initial hypothesis increased. Both changes were in the direction predicted by the results of Bruner *et al.*

TABLE I

PERFORMANCE AS A FUNCTION OF SIZE OF INITIAL HYPOTHESIS

	Number of attributes included in initial hypothesis					
	1	2	3	4	5	6
Number of Problems	38	46	30	17	5	7
Mean trials to solution	10.11	9.39	9.60	9.06	8.25	7.00
% solutions attained seven trials	50	59	53	53	75	100

There are four types of instances which *S* might encounter in a problem of concept-attainment: (*PC*) the positive confirming instance with respect to the hypothesis currently held by *S*; (*PI*) the positive infirming instance not in accord with *S*'s current hypothesis; (*NC*) the negative confirming instance; and (*NI*) the negative infirming instance. Bruner *et al.* have shown that, regardless of *S*'s strategy, it is of utmost importance for efficient solving of problems that *PI* instances be handled properly. Supporting evidence for this conclusion was derived from the data of the present experiment. First, among the problems attacked by the wholist strategy, 51% of the total instances encountered were of the *PI* variety. Of these 85% were handled correctly, and of those *S*s who handled all such instances correctly 100% solved their problems within seven trials. Thirty-three percentage of the instances encountered by the partists were *PI,* and of these 45% were handled correctly. Note that this percentage is far below that of the wholist and attests to the inferiority of the partist approach. Of those partists who handled all *PI* contingencies correctly, 85% solved their problems within the seven trials; of those *S*s who did not handle *PI* instances correctly, only 27% solved within seven trials. This relationship can be further documented by the product-moment correlation between trials to solution and percentage of *PI* instances handled correctly ($r = 0.374$, 130 *df*, $p < 0.01$).

Amount and type of stimulus-variation, complexity and method of stimulus-presentation affected the performance of partists and wholists in a way which was consistent with the influence of these variables on overall trials to solution. As amount of stimulus-variation shifted in value from one through three attributes, the percentage of partists solving with minimal information dropped, taking the values 61%, 58% and 40%, respectively. Corresponding percentages for the wholists, who were few in number, were 100%, 100%, and 75%. A more marked effect would be expected in the case of the partists on the grounds that such a strategy may permit *S* to overlook much of the information presented by large changes in the stimulus during the early trials of the problem. Since certain attributes were not considered important initially by the partists, and since these 'unimportant' attributes may be shown relevant or irrelevant to the concept during the first few trials, certain indispensible information may be bypassed until later instances. Such cannot be the case with the ideal wholist approach which considers all information potentially relevant until proved otherwise. No matter how large the variation of stimuli, overlooking the memory capacity of *S* momentarily, no available information goes unconsidered. Sixty percentage of the partists under the condition of shifting focus as opposed to 46% under fixed focus solved within seven trials. Problems with only one relevant attribute were solved within seven trials by partists in 76% of the cases, whereas the two- and three-attribute problems were solved by only 52% and 31%, respectively. Finally, it was observed that 39% of problems initiated with the partist approach were solved within seven trials when only one stimulus at a time was available to *S*. Sixty-eight percentage solutions were obtained when the instances were arrayed. These data confirm the conclusion of Bruner *et al.* that any condition which increases the amount of information *S* must contend with at any one time will reduce the effectiveness of a strategy, such as the partist, which requires relatively sizeable feats of memory itself.

SUMMARY

The purpose of this experiment was to study the influence of certain factors on the plans or strategies and the over-all efficiency of performance of *S*s in attaining concepts. The design of the experiment incorporated four major variables: amount of trial-to-trial variation of stimuli, type of variation of stimuli, number of attributes relevant to a concept, and method of presenting stimuli. The basic design was a 2 × 2 × 2 factorial within which a 3 × 3 Greco-Latin square was used to determine the amount and type of stimulus-variation for each of three problems given 48 *S*s.

The results of the study were: (1) Amount of stimulus-variation did not significantly increase mean trials to solution of the problem although the percentage of *S*s solving the problems with minimal information decreased. (2) Type of stimulus-variation was statistically significant as was its interaction with method of presentation. (3) When *S* was forced to rely on memory of previous instances, performance was significantly impeded. (4) Increases in complexity of the task produced increases in difficulty. (5) The two measures of response, trials to solution and percentage of *S*s solving with minimal information, yielded similar and consistent results. (6) In the main, attempts to identify *S*s using the partist and the wholist strategies described by Bruner *et al.* were successful, and results were in accord with those reported earlier. On all measures the partists were poorer performers, although this result has to be interpreted in light of the fact that the wholist strategy was used in only 9% of the problems.

Results were interpreted within the model formulated by Bruner *et al.* and served to supplement available data related to the use of strategies in the formation of concepts.

APPROXIMATIONS TO ENGLISH: SOME COMMENTS ON THE METHOD

By E. B. COLEMAN, Sul Ross State College

Miller and Selfridge plotted approximations to English against percentage of words recalled and got a negatively accelerated curve that became horizontal at about a fifth-order approximation. From this they concluded that a fifth-order approximation "is as easily recalled as a passage lifted from a novel . . . meaningful material is easy to learn, not because it is meaningful *per se,* but because it preserves the short range associations that are familiar to the *S*s."[1] Twelve years have gone by and the authors themselves have modified their conclusions,[2] but the shape of the curve and the original conclusions are still important in the literature. The curves and conclusions are quoted in general texts enjoying wide circulation,[3] and the results have been replicated many times.

It is true that Marks and Jack reported conflicting results: they found that a significantly higher percentage of the prose-selections was recalled.[4] They used, however, a different measure, *i.e.* the number of words recalled correctly in sequence up to the first error—the immediate memory span.

Miller and Selfridge had ignored both order and intrusions, merely counting total number of words correctly recalled. When this measure was used, many investigators have replicated the results of Miller and Selfridge: their curves became horizontal at about fifth-order approximations.[5] In fact, the *S*s of one investigator (Sharp) made slightly lower scores on prose than on higher-order approximations.

---

* Received for publication September 21, 1961.

[1] G. A. Miller and J. A. Selfridge, Verbal context and the recall of meaningful material, this JOURNAL, 63, 1950, 176-185.

[2] At any rate, the conclusion of this paper is little more than a paraphrase of Miller's notion of recoding (G. A. Miller, The magic number seven; plus or minus two: Some limits on our capacity for processing information, *Psychol. Rev.,* 63, 1956, 81-97). Also see Miller, Human memory and the storage of information, *I. R. E. Transactions on Information Theory,* 1956, 2, 135.

[3] James Deese, *The Psychology of Learning,* 2nd ed., 1958, 177-179, E. R. Hilgard, *Theories of Learning,* 2nd ed., 1956, 381-382, Miller, *Language and Communication,* 1951, 212-213.

[4] M. R. Marks and Ollie Jack, Verbal context and memory span for meaningful material, this JOURNAL, 65, 1952, 298-300.

[5] Three of the most recent replications were: Leo Postman and P. A. Adams, Studies in incidental learning: VIII. The effects of contextual determination, *J. exp. Psychol.,* 59, 1960, 153-164; Patricia Richardson and J. F. Voss, Replication report: Verbal context and the recall of meaningful material, *ibid.,* 60, 1960, 417-418; H. C. Sharp, Effect of contextual constraint upon recall of verbal passages, this JOURNAL, 71, 1958, 568-572.

Sharp's unexpected finding is not unique. Using a different measure, 'cloz' scores, Deese repeated Sharp's results.[6] After fifth or sixth order, 'cloz' scores began to decrease. Thus, for two different measures and for two different samples of approximations to English, we find that after about fifth or sixth order, higher-order approximations begin to diverge from *S*'s verbal habits.

It is really quite surprising to find that the only associations important to recall are those extending over sequences of less than six words. Language can be analyzed into units larger than words—clauses and paragraphs for instance. Surely the long-range, high-order associations between these larger units are important in recall. The mere flattening of the curve is surprising; but when both Deese and Sharp find the curve turning back downward, the implications become unreasonable enough to bring the experimental technique under suspicion. The grammatical constructions in higher-order approximations have always seemed to be more far-fetched and implausible than the constructions in lower orders. This personal impression, if it were corroborated by more objective evidence, would explain the flattening of the curve. As a matter of fact, it was corroborated by the following scaling procedure.

EXPERIMENTAL I: SCALING PROCEDURE

(*A*) *Ranking approximations as to grammaticalness: Judges.* Twelve linguists ranked 12 sets of approximations to English as to their grammaticalness. All the judges were attending the 1961 Linguistic Institute meeting at the University of Texas. Six held doctoral degrees in either English or linguistics; six were graduate students in linguistics. None of them was familiar with approximations to English.

Materials. One set of approximations was drawn from the set collected by Sharp.[7] It consisted of a second-, a third-, and a fifth-, and a ninth-order approximation. A zero-order approximation was added by sampling at random from *The Teacher's Word Book*,[8] and a first-order was added by sampling randomly from Sharp's other approximations. The first 15 words of each approximation were typed on an individual card.

The other 11 sets were similar. Three were drawn from Miller and Selfridge's collection, two from Moray and Taylor's,[9] and six from Deese's.[10] The individual sequence on each card was 15 words long, and all sets were supplied with a zero- and a first-order approximation if they did not already contain them. Thus there were 12 scales represented by 12 packets of cards.

---

[6] Deese, From the isolated unit to connected discourse, in C. N. Cofer (ed), *Verbal Learning and Verbal Behavior*, 1961, 11-31.

[7] Sharp, *op. cit.*, 568.

[8] E. L. Thorndike and Irving Lorge, *The Teacher's Workbook of 30,000 Words*, 1944.

[9] Neville Moray and Ann Taylor, The effect of redundancy in shadowing one of two dichotic messages, *Language and Speech*, 1, 1958, 102-109.

[10] Deese, *op. cit.*, 177.

Procedure. A judge was given a packet of cards and asked to rank them as to grammaticalness. He was instructed to use anything he knew about English grammar as a basis of his judgment.

Results and discussion. In Fig. 1, mean rank assigned by the judges is plotted against order of approximation.[11] The curve shows clearly that the linguists thought that the higher-order approximations deviated from English grammatical rules just as much as did the fifth-order approximations. As order of approximation becomes higher, apparently the grammatical constructions gradually become more awkward and implausible; that is, the subjects do not agree so well with their verbs, nor the verbs with their direct objects, nor the modifiers with the words they modify, etc.

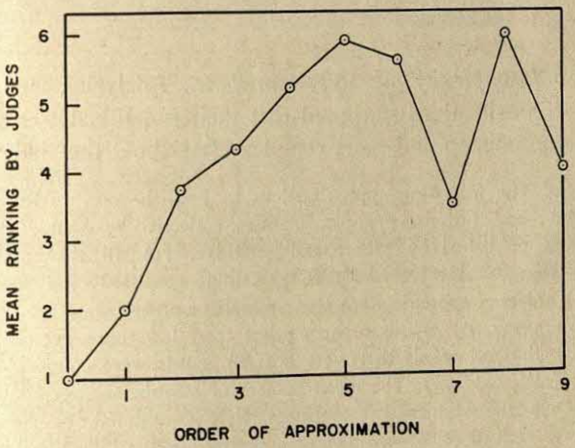


FIG. 1. MEAN JUDGES' RANKING OF GRAMMATICALNESS PLOTTED AGAINST ORDER OF APPROXIMATION TO ENGLISH

A gradually increasing deviation from common English is not hard to explain when the technique for generating approximations is considered. When the person who originally generated an approximation was faced with a short sequence (*humble because,* for instance), he could incorporate it into a short, simple sentence. On the other hand, when he was faced with a long sequence (*humble because they have no electricity*), sometimes he had to resort to a complex and far-fetched sentence to incorporate it. The effect was multiplied because he was not allowed to use punctuation. Thus, as order of approximation increased, the sentences from which the approximations were derived steadily diverged from common English.

---

[11] This curve contradicts a statement by Chomsky, and may be of some incidental interest to linguists. N. A. Chomsky states "there appears to be no particular relation between order of approximation and grammaticalness." (N. A. Chomsky, *Syntactic Structures,* 1957, 17.)

The implications of Fig. 1 are obvious: its curve is almost identical to the usual curve representing recall of approximations to English. This suggests that the previous recall curves became horizontal because increasing grammatical awkwardness of some sort was counterbalancing the effect of long-range associations.

(B) Word-complexity of approximations. Analyzing an extensive sample of approximations suggested that their individual words may become more uncommon and more complex after about the sixth-order.

The sample. The following approximations to English were analyzed: (a) the second-, third-, and fifth-order approximations collected by Sharp, (b) all the approximations published by Miller and Selfridge, (c) all the approximations collected by Deese. Table I shows that the sample consisted of 4100 words and represented 9 orders of approximation plus a sample of prose.

Uncommon words. All words occurring less than 100 times per million words were counted, thus all words that were not AA words were counted.[12]

Morphological complexity. The mean number of morphemes per 100 words was computed. Only the segmental morphemes were considered, of course, and a morpheme was defined as a word-root or a derivational or inflectional affix except that suppletive and replacive inflections were not counted. For instance, *unearned* was counted as three morphemes: *un, earn, ed;* but *were* was counted as a single morpheme.

Clearly this is the same sort of complexity that Flesch is measuring less directly by counting syllables.[13] As a matter of fact, Flesch originally counted affixes,[14] and this is almost identical to computing the mean number of morphemes per word.

Syllables per 100 words. Because it is so simple and reliable, the number of syllables per 100 words was also counted as a second index of morphological complexity.

Results. Table I shows that the first-, third-, fourth-, fifth-, sixth-, and seventh-order approximations are about equally complex, but above the seventh-order the approximations may increase rapidly in word-complexity. The eighth-order passages jumped to 1.48 syllables per word and to 18% not AA. The indices for the prose are even higher. The prose has indices equivalent to material at the high-school level (its Flesch Reading

---

[12] Thorndike and Lorge, *op. cit.*, ix.

[13] Rudolf Flesch, *The Art of Readable Writing*, 1949, 214.

[14] Flesch, *Marks of Readable Style*, 1942.

Ease Score was 58 or 'difficult'), but the approximations have indices roughly equivalent to third-grade readers.

Two other findings in Table I deserve comment though they are not directly related to the argument of this paper. First, zero-order approximations are unusually complex because a list drawn at random from a dictionary contains a high proportion of morphologically complex words such as *unbeliev-able*. Secondly, the second-order approximations are unusually simple because *is* and its inflected forms occur frequently, resulting in cycles such as *"you are you are you."*

TABLE I

NUMBER OF SYLLABLES, MORPHEMES, AND DIFFICULT WORDS PER 100 WORDS FOR ORDERS OF APPROXIMATION TO ENGLISH

	Order of approximation									
	0	1	2	3	4	5	6	7	8	Text
Words in sample	110	610	620	620	610	620	500	110	500	110
Syllables per 100 words	244	135	128	137	134	137	136	135	148	157
Morphemes per 100 words	196	120	114	119	122	119	121	119	128	132
Difficult words per 100	96	12	5	13	12	13	14	13	18	25

To summarize, the analysis agreed with the ranking by linguists: they both suggested that the curves of recall in previous investigations became horizontal at fifth-order approximations because increasing complexity was counterbalancing the effect of long-range associations. If the prose and all approximations were matched for word-complexity, the advantage of long-range associations might be demonstrated.

EXPERIMENT II: RECALL OF APPROXIMATIONS TO ENGLISH

In the following two experiments (A and B) *S*s attempted to recall orders of approximation to English.

These experiments differ from previous, similar studies in two respects: (1) Insofar as possible, word-complexity was held constant. All passages were matched for frequency of occurrence of words and for syllabic length. (2) Recall was scored in long sequences. Besides being scored in correct single words, recall was scored in correct two-word sequences, correct three-word sequences, and on up to correct 17-word sequences. This dimension of scoring systems will relate the findings of Marks and Jack with those of the other investigators. Essentially, Marks and Jack scored in long sequences: they used number of words recalled correctly up to the first error. By scoring in terms of single words, the other investigators may have favored the short-range associations over the long-range ones.

Experiment IIA: Subjects. A class of 32 students in introductory psychology served as *S*s.

Materials. Approximations collected by Deese that were 50 words long were cut to give 30-word passages. The passages varied greatly in difficulty, but they were grouped in four sets so that within each set the passages were matched closely for syllabic length and percentage of AA words. Two of the sets were first-, third-, and fifth-order approximations; the other two were second-, fourth-, and sixth-order approximations.

A passage was drawn from modern prose to match each set, making four passages in each set. Thus each set was a scale that gradually approached the statistical structure of English, and the passages in each scale were closely matched in syllabic length and AA words. In syllabic length per word, the prose averaged 1.37 and the approximations 1.38. In percentage of words not AA, the prose averaged 15.2% and the approximations 15.0%.

Presentation. The approximations of a single set were presented in random order to the entire class. *E* selected one set, shuffled it, took the top passage, and read it to the class in a monotone. After he finished, the *S*s wrote down all they recalled. They were instructed to write the words in order if possible, but to write all words they remembered even if they forgot the order. When all members of the class had finished, *E* read the second passage of the set, and so on through the set. Then he selected a second set, shuffled it, and so on through all four sets.

Scoring. The *S*s were scored for each single word correctly recalled regardless of order, every correct two-word sequence, and so on up to correct 17-word sequences. (For instance, ABCDEF has the following three-item sequences: ABC, BCD, CDE, DEF.)

Experiment IIB. Experiment IIB also tested the recall of orders of approximation. Its procedures differed from IIA in the following respects: (a) three different lengths were examined: 50-, 100-, and 150-word passages; (b) The sample of approximations was increased: there were six sets; and (c) The sample of *S*s was decreased to 12. The previous experiment had suggested that most of the random error was due to the linguistic sample. (d) The *S*s read the passages, which were typed on cards in solid capitals with no punctuation. They were given 0.5 sec. per word and instructed to read the passage and memorize as much as possible. Otherwise the presentation resembled that in Experiment IIA: orders of approximation were presented in random order, and immediately after reading a passage, *S* wrote it down.

Results. In Fig. 2, the mean number of items correct is plotted against order of approximation for the 30-word sequences of Experiment IIA. The curves for the longer sequences of Experiment IIB are almost identical. (In Fig. 2, note that adjacent orders of approximation were averaged together so that each point would represent four passages as did the point representing recall for prose.) All 32 *S*s made higher scores for all methods of scoring on prose than on fifth-order approximations. All the *S*s also made higher scores on prose than on the sixth-order approximation. Such a difference is significant, of course, at one over 2^{32}.

The curve between 5½-order and prose is staggered to remind the reader that the abscissa should actually be extended far to the right. Of course, if it were extended to its rightful position, the curves would become negatively accelerated. This seems reasonable because a twenty-

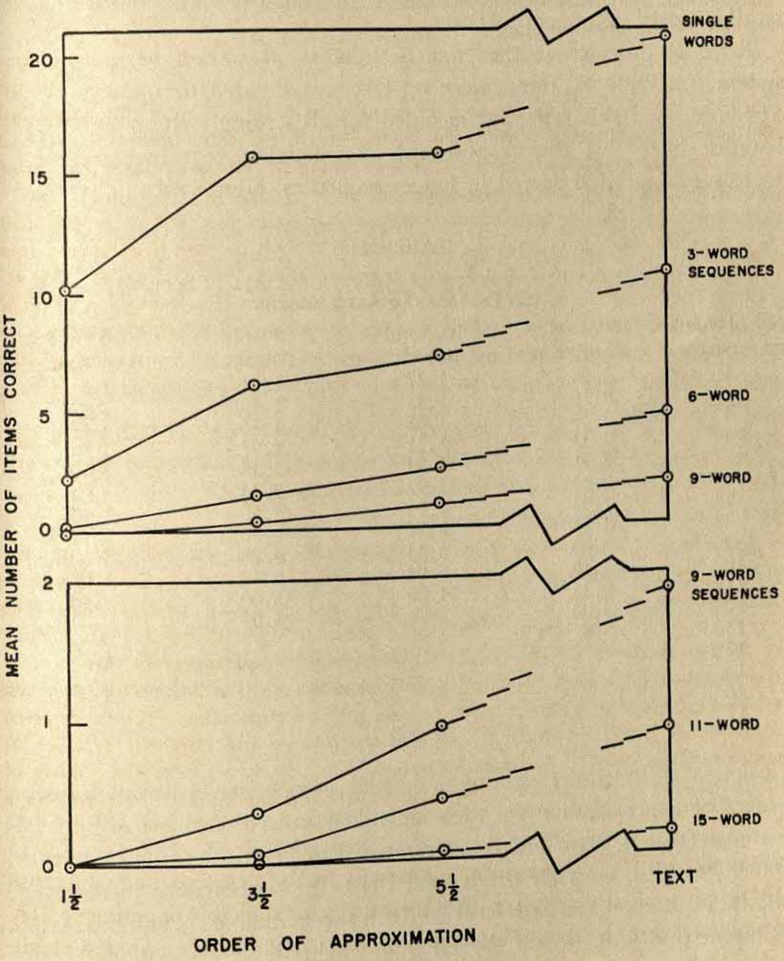


FIG. 2. MEAN NUMBER OF ITEMS CORRECT PLOTTED AGAINST ORDER OF APPROXIMATION

ninth-order approximation should be remembered almost as well as the prose (which might be called a thirtieth-order approximation).

The rather obvious interaction between order of approximation and method of scoring is also significant. A plot similar to Fig. 2 was made for each *S*. For 23 *Ss*, the curves became more and more positively accelerated as the sequences scored became longer; for eight *S*s, the curves became less positively accelerated, and the final *S*'s curves were indeterminate. Thus the interaction is significant beyond the 0.01 level by a binomial test: the advantages of higher-order approximations increase as recall is scored in longer sequences.

Table II provides another way to interpret the effect of the scoring system. In Table II, mean score on text was divided by mean score on 5½-order, and this ratio was recorded for all scoring systems and for all lengths. Note that the proportional advantage of text over 5½-order increases as recall is scored in longer sequences. Significance of this conclusion was tested by making a similar comparison between adjacent orders of approximation for each individual and for each set of approximations (that is, sixth order was compared to fifth order, fifth to fourth, fourth to third, etc.) By the binomial test, the proportional increase was significant beyond the 0.01 level for both the sample of *S*s and the sample of approximations. Again the effect of higher-level associations became more and more apparent as recall was scored in longer sequences.

TABLE II

SCORE ON TEXT DIVIDED BY SCORE ON 5½-ORDER FOR ALL 17 SCORING SYSTEMS AND PASSAGES OF ALL LENGTHS

No. of words in sequence	150-word passage	100-word passage	50-word passage	30-word passage
1	2.08	1.27	1.21	1.31
2	2.16	1.43	1.49	1.47
3	2.29	1.65	1.76	1.49
4	2.43	1.71	1.97	1.62
5	2.81	1.82	2.28	1.70
6	2.92	1.97	2.45	1.82
7	2.93	2.16	2.80	1.97
8	3.33	2.34	3.36	1.98
9	4.06	3.43	4.28	2.02
10	5.10	5.06	6.17	1.98
11	9.25	11.20	15.00	2.20
12	28.00	∞	25.00	2.42
13	∞		∞	2.95
14				4.00
15				5.65
16				13.00
17				∞

DISCUSSION AND CONCLUSIONS

In the experiments of earlier investigators, recall for prose and higher-order approximations must have been depressed because these passages were more complex and grammatically awkward than the lower-order ones. To an extent, their curves became horizontal because increasing complexity was counterbalancing the effect of long-range constraints—the constraints that package phrases, clauses, and paragraphs into familiar orders.

The present study demonstrated that when passages were matched in word-complexity higher percentages were recalled as the passages approached the statistical structure of language, and there was no leveling off at fifth-order approximations. Furthermore, as recall was scored in longer and longer sequences, the effect of long-range constraints became more and more apparent.

In recalling approximations to English, *S* is aided by positive transfer from his verbal habits, mainly from his habits that govern the order and co öccurrence of verbal elements. He finds adjacent elements in a familiar order, or at least the constraints upon order and co öccurrence resemble the constraints of his own habits. In zero-order approximations, the letters are ordered as familiar words. At second- and third-order approximations, additional constraints package the words into more or less familiar phrases: non-human nouns, for instance, will seldom be modified by adjectives such as *erudite*. In higher-order approximations, additional constraints package the elements into still more familiar orders: a clause (or even a paragraph) about organic farming, for instance, will seldom be adjacent to one about submarine warfare.

Clearly *S* has less to learn if the order of the list is already familiar, and the effect of this positive transfer from his extra-experimental ordering habits will become more and more apparent as the scoring system increases its emphasis upon order.

SUMMARY

A ranking by linguists and an analysis of word-complexity both suggested that the technique used to generate approximations to English produce higher-order approximations that are more complex and grammatically deviant than lower-order ones. As the person who originally generates the approximations is faced with longer sequences to be incorporated into a sentence, he apparently resorts to more and more complex and far-fetched sentences.

The experiment of Miller and Selfridge was repeated with two changes: (a) the prose and all the approximations were matched in syllabic length and in word-frequency; (b) recall was scored in correct two-word, correct three-word, and on up to correct 17-word sequences. Prose was recalled significantly better than the higher-order approximations. Also the advantages for the higher-order approximations became greater as recall was scored in longer sequences.

COLOR-NAME AS A FUNCTION OF WAVE-LENGTH

By ALEEZA CERF BEARE, The Johns Hopkins University

This study is concerned with the way in which color-names vary with wavelength at each of two levels of luminance. Instructions restrict responses to the basic names ususally considered as applying to the spectrum —*red, orange, yellow, green, blue,* and *violet*—a six-category system for which Graham has provided a set of hypothetical curves.[1] One and the same set of determinations provides the basic color-name for a given wavelength (R_A), the judgment-time for this response (R_J), and, after the response in terms of the restricted set of categories, a qualifying response (R_B). The qualifying response is given after the exposure is terminated; it is optional and unrestricted, intended to permit 'better description' of the chromatic flux than is possible with the six categories.

Although little studied by psychologists, color-naming has been examined by anthropologists, such as Rivers[2] and, more recently, Ray,[3] in comparative linguistics,[4] and in studies of communication.[5] None of these studies has, however, revealed anything about the relation of the basic color-names to wavelength, data that are needed to provide a standard reference in the study of visual perception. In addition to establishing these color-naming functions, it will be of considerable interest to examine the relation of such variables as judgment-time and frequency of other names to the basic functions and, specifically, whether a relationship can be demonstrated between judgment-time and the use of qualifying terms. On the basis of existing data from other perceptual functions, these two measures may be expected to show a positive relationship, rising

---

Received for publication September 28, 1961.

* This paper is based on a doctral dissertation submitted to Columbia University. The research was supported by the Office of Naval Research. The author is indebted to Professor Clarence H. Graham, under whose supervision the research was conducted, and to Drs. Yun Hsia and William J. McGill for their assistance.

[1] C. H. Graham, Sensation and perception in an objective psychology, *Psychol. Rev.*, 65, 1958, 65-77.

[2] W. H. R. Rivers, Observations of the senses of the Todas, *Brit. J. Psychol.*, 1, 1905, 326-339.

[3] V. F. Ray, Human color perception and behavioral response, *Trans. N.Y. Acad. Sci.*, 16, 1953, 98-104.

[4] E. H. Lenneberg and J. M. Roberts, The language of experience, *Int. J. Amer. Linguistics*, 22, 1956, (No. 2), 1-33.

[5] R. M. Halsey, Identification of signal lights, *J. opt. Soc. Am.*, 49, 1959, 45-55, 167-169.

in the region of subjective equality between any two categories, as well as at points where naming changes rapidly, or where some perceptual ambiguity exists.[6] It must be borne in mind, however, that judgment-time may rise also as a function of variables other than ambiguity or lack of confidence on the part of *O*. According to Hick, there is as yet no independent measure of reaction-time (or judgment-time) with regard to individual stimulus-categories.[7]

Although little or no experimental evidence is available, color-naming tendencies may be thought of as formed during childhood, along with the general development of the linguistic apparatus. Adult *O*s serving in a color-naming experiment of the sort reported here make use of a 'built-in' and 'readily-available' response-system.[8] It is assumed that the 'built-in' linguistic apparatus is sufficiently uniform to provide frequency distributions which represent valid color-naming functions for the six categories employed.

METHOD

Apparatus. A Bausch & Lomb grating monochromator, having as a source a tungsten ribbon-filament bulb rated at 6 v., 18 a., AC, was connected to a Variac which controlled the voltage of the light-source. On the basis of the *cff*-procedure (see below), a 6-v. source was selected for the monochromator's high-intensity setting (35 c.p.s.) and a 5-v. source for the low-intensity setting (20 c.p.s.). Measurement with a Macbeth illuminometer of white light at the same *cff* values gave: low intensity-= 3100 mL, high intensity = 7410 mL.

The monochromator joined a light-tight, three-lens optical system providing a Maxwellian view at the eye, and a filter-system with neutral-density filters, a pair of polaroids, and three color-filters. The stimulus-field subtended a visual angle of 2°. A solenoid-mounted, vane shutter controlled exposures; it was operated by a DPST knife-switch which also controlled a timer, providing cumulative judgment-time in 1/100 sec.

A DC-motor-powered Variac speed-control drove a disk of 5 in. radius with two alternate 90° cut-out sectors to provide flicker. Motor and speed-control were so mounted that their vibration would not affect the rest of the apparatus or *O*'s cubicle. This cubicle, which remained open during experimental sessions, was painted flat black and contained the eye-piece and an adjustable chin-rest. The eye-piece (with artificial pupil) was surrounded by a sheet of flat white drawing paper, having a luminance of 6.05 mL, which was illuminated by an overhead 200-w. tungsten bulb (of color temperature 2700° K), kept burning throughout the experiment.

---

[6] J. P. Guilford, *Psychometric Methods*, 2nd ed., 1954, 302-307.

[7] W. E. Hick, On the rate of gain of information, *Quart. J. exp. Psychol.*, 4, 1952, 11-26.

[8] Alphonse Chapanis and R. M. Halsey, Absolute judgments of spectrum colors, *J. Psychol.*, 42, 1956, 99-193.

Observers. Ten subjects, four women and six men, ranging in age from 18-27 yr. participated in the experiment. They were graduate or undergraduate students at Columbia University. All had been tested for color-normality by the Ishihara Test (13th Edition) and the Hecht Schlaer Anomaloscope, and all claimed English as their native tongue.

Procedure. The *cff*-sessions, which preceded the naming sessions, were conducted to equate luminance throughout the spectrum, on the assumption that an equal fusion-rate represents equal luminance. For low intensity, flicker was held constant at 20 c.p.s. and for high intensity at 35 c.p.s. Testing proceeded from 400 mμ, in steps of 7 mμ, as far into the red end as was possible, while luminance was adjusted by means of the filter-system for each spectral point until *O* reported fusion. An exception to the 7-mμ step had to be made in the narrow yellow band where the steps were decreased to 4 mμ. *Cff*-determinations were made with an ascending order of the method of limits, and each session lasted about 1 hr. A period of 10-20 sec. separated determinations, during which time *O* was instructed to look at the white paper surrounding the eye-piece.

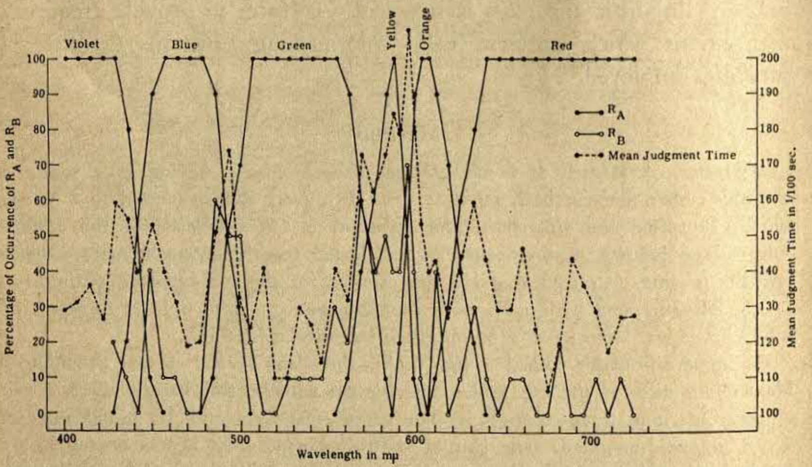


FIG. 1. FREQUENCY-DISTRIBUTIONS FOR SIX COLOR-NAMES, DISTRIBUTION OF QUALIFYING RESPONSES, AND OF MEAN JUDGMENT-TIME AT LOW INTENSITY

Before each naming session, *O*'s relevant *cff* data were rearranged according to a table of random numbers, with the result that each naming session proceeded according to a unique random order. The filter-conditions determined by *cff* prevailed for each naming exposure, and testing followed the method of absolute judgments. The *O* was instructed to look into the eye-piece at the 'ready' signal and then, as rapidly as possible, name the color exposed, using one of the following names: *violet, blue, green, yellow, orange, red.* As soon as he had given his response (R_A), the exposure was terminated, and *O* then could qualify his judgment however he wished (R_B). Response R_A, R_B if given, and R_J, the judgment-time for R_A in 1/100 sec., were recorded for each exposure. Between successive exposures, *O* was instructed to look at the paper surrounding the eye-piece. Naming sessions lasted from 1–2 hr. and were interrupted by a 10-min. rest-period halfway through the determinations.

RESULTS

The results are summarized in Fig. I for the low intensity and in Fig. 2 for the high intensity. Each set of curves shows a discrete frequency distribution for each of the six color-names, and one continuous curve each for R_B and mean R_J, all as a function of wavelength. The only difference between high and low intensity here is that for high intensity, no 100% point is obtained for violet, and the 100% range for red is shorter than that of the low intensity. This difference is in accord with what is generally known about the luminosity-curve.[9]

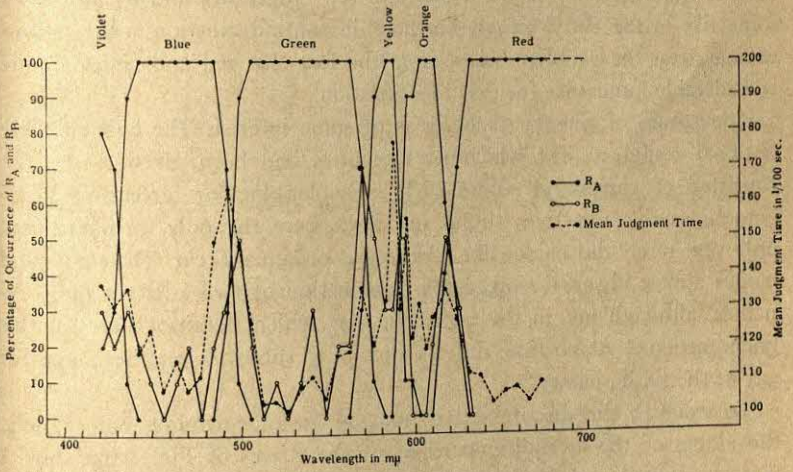


Fig. 2. Frequency-Distributions for Six Color-Names, Distribution of Qualifying Responses, and of Mean Judgment-Time at High Intensity

Both illustrations demonstrate that the ascending portion of each R_A-curve is symmetrical with the descending portion of the preceding curve—that the *PSE* between categories occurs in every case at the 50% point. Apparently, only two 'adjacent' names ever competed for the R_A-response and no overlap from a third name entered the choice, even for the region from 550 610 mμ, where the names *green*, *yellow*, and *orange* occur in rapid succession. This finding presents the only deviation from the color naming functions predicted by Graham, who placed the crossover point from blue to green at a point above 50% on the assumption that yellow and violet responses might make some contribution.[10]

---

[9] Graham, Color theory, in S. Koch (ed.), *Psychology: A Study of a Science, 1,* 1959, 145-288.

The assumed relationship between R_B and R_J is suggested by Fig. 1 and Fig. 2. The peaks of both curves actually cluster around the *PSE*'s and generally tend to remain low for each range of $R_A = 100\%$.[11] The additional fact that certain wavelength-bands seem to impose specific conditions on namability emerges from the data. Identification appears difficult where a rapid change of name occurs, *i.e.* where R_A comes to a peak at 100% rather than extend there over a range of wavelengths, notably in the yellow-orange area, where R_B and R_J tend to remain high even at $R_A = 100\%$. In the violet and red ranges, R_B and R_J also show some rise in the 100% range, but here, in contradistinction to the yellow-orange area, the problem arises from the fact that red and violet require considerable luminance for easy identification.

The nature of the R_B-responses is of some interest. The bulk of these optional responses, for which no categories had been given to the *O*s, consisted of compound names such as 'violet-blue' or 'greenish-yellow.' Sometimes the adjectives 'light' or 'dark' were the only qualifiers, and only very rarely did names like 'chartreuse' or 'aqua' occur. The compound names almost always involved the combination of two adjacent category-names, although not in the systematic way which Bouma claims for this phenomenon.[12] At no time did *O*s attempt to substitute another name for any of the six R_A names.

To examine the discriminative function in color-naming more closely, the slopes of the ascending portions of the curves of Fig. 1 and Fig. 2 were computed and are plotted as a function of wavelength (the 50% points were used as plotting points) in Fig. 3. It appears that the naming slopes for all six category-names are fairly uniform except for the slope denoting the rise from yellow to orange, which is of a higher order of magnitude. In Fig. 1 and Fig. 2, the range separating 100% yellow from 100% orange is the shortest, and it may be reasonable to expect that the shorter the range the steeper the curve must be to reach its maximum; however, since rank-order correlations between range in mμ and slope showed a trend towards significance only for the data of Fig. 1, the explanation of this phenomenon must be sought elsewhere.

---

[10] Graham, *op. cit.*, 65, 1958, 65-77.

[11] The term 'range' as used here always refers to a specified segment of the physical-wavelength continuum as measured in mμ.

[12] P. J. Bouma, *Physical Aspects of Colour*, 1947, 16.

To test the color-naming data against those of hue-discrimination, an attempt was made to fit the former to the hue-discrimination curve given by Judd.[13] The 100% R_A-ranges in mμ, from Fig. 1 and Fig. 2, were reduced by a factor of 20 and the results of this transformation (as plotting points the 50% *loci* were again chosen) are shown in Fig. 4. The arrow at each end of the range (or Δλ) curves indicates that the Δλ value calculated for color-name ranges must be greater than the one shown (the frequency-curves having indeterminate limits at the ends of the spectrum), even though no greater value could be obtained with the

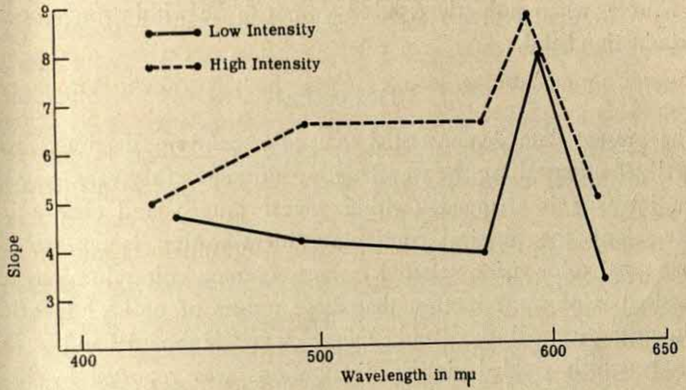


FIG. 3. SLOPES OF COLOR-NAMING CURVES AT HIGH AND LOW INTENSITIES

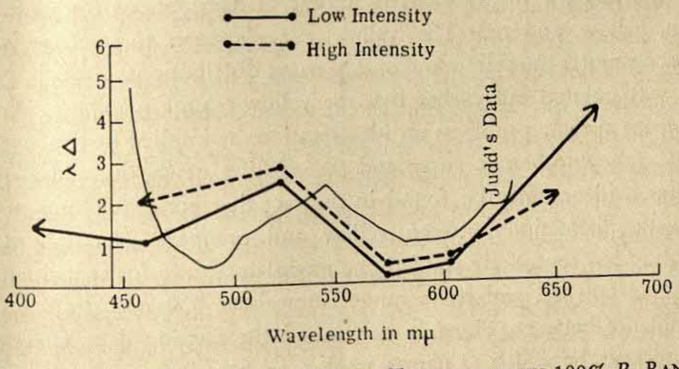


FIG. 4. JUDD'S HUE-DISCRIMINATION CURVE TOGETHER WITH 100% R_A-RANGES AT HIGH AND LOW INTENSITIES

conditions under which the present data were collected. Thus, what fit there is between the two functions is tenuous, and it is possible that the configuration of the naming curve will change if greater values are obtained in the red and violet ends, or with changes in instructions with regard to categories of judgment. For the present data, however, the range-curve, when properly reduced, comes to lie within the general hue-discrimination band.

---

[13] D. B. Judd, Chromaticity sensibility to stimulus differences, *J. opt. Soc. Am.*, 22, 1932, 72-108.

DISCUSSION

The present data demonstrated that color-naming functions form discrete distributions along the wavelength-continuum; they vary only slightly when intensity is changed (within foveal range) and clearly indicate where, along the physical continuum, identification is optimal for *O*s with normal color-vision selected from a common cultural-linguistic background. It is of some interest that these ranges of optimal identification are in agreement with the data of Dimmick and Hubbard[14] and of Purdy.[15] A study which yielded similar distributions was reported by Ekman,[16] who applied a factor analysis to ratings of spectral colors for degree of 'qualitative similarity.' Five factors were thus obtained which, when plotted over the baseline of wavelength, appear as distributions for violet, blue, green, yellow, and red. The 'yellow' curve appears to combine into one what constitutes the yellow and orange name-distributions.

It was pointed out earlier that the yellow-orange naming area seemed to impose specific problems on identification, as implied by the steep slope of the yellow-to-orange curve and the peaking of the R_B and R_J curves. Further evidence may be found in the fact that not all cultures make the linguistic distinction between yellow and orange.[17] Neurophysiological recording has repeatedly yielded one modulator-curve for red which shows a peak at 600 mμ and has a range from about 530–650 mμ.[18] From the resemblance between Ekman's curves and the naming data, it seems that the processes by which *O* names a color, or rates it for 'subjective similarity' with another, are governed by the same or related variables. The difference between the two response-configurations must be assumed to lie in the different instructions inherent in the differing psychophysical methods employed, and the question which must be asked is: What naming distributions could have been obtained had *O*s been asked to supply their own naming categories (different method) or if more than six, or fewer categories had been used within the method of absolute judgments?

---

[14] F. L. Dimmick and M. R. Hubbard, The spectral location of psychologically unique yellow, green, and blue, this JOURNAL, 52, 1939, 242-251.

[15] D. McL. Purdy, Spectral hue as a function of intensity, this JOURNAL, 43, 1931, 541-559.

[16] Gösta Ekman, Dimensions of color vision, *J. Psychol.*, 38, 1954, 467-474.

[17] Lenneberg and Roberts, *op. cit.*, 1956, 30.

[18] Ragnar Granit, *Receptors and Sensory Perception*, 1955, 137.

The direct influence which instructions are likely to exert on the response-configuration is probably greater than is commonly realized. It has correctly been argued by Halsey and Chapanis that the limiting of categories might lead to the grouping of a number of wavelengths which with different instructions, may have been distinguished perfectly.[19] Alluisi has shown that number of response-categories seems so to interact with number of stimulus-categories that, when the number of the former is less than the number of the latter, the amount of information transmitted appears to be less than when the number of response-categories is equal to or greater than the number of stimulus-categories.[20] Both statements imply a lack of distinction between a discrimination- and an identification-task and ignore the fact that the retina's capacity for hue-discrimination is not matched by man's linguistic equipment. The latter provides for the grouping of spectral stimuli (though it does not exclude the possibility of identifying the colors of name-bands in the spectrum). The kind of rating experiment employed by Ekman is much more like a discrimination task (*O* responds in terms of 'more' or 'less') than the naming experiment and yet is again limited by the number of stimuli compared with one another.[21] With a larger number of comparisons, separate curves for yellow and orange might be obtained. The possibility must be considered, however, that the yellow-orange dichotomy is an artificial one, and that, if either category were omitted from the instructions, the fourth of the resulting set of naming curves would have a 100% R_A-range extending from about 586-606 mμ and would exhibit a slope which would be of the same order of magnitude as are those of the other naming curves. The name which would best represent the yellow-orange area must, of course, be determined experimentally.

If the number of naming curves were reduced from six to five by the process just outlined, it is likely that the resemblance of the curve representing 100% R_A-ranges to the hue-discrimination curve would become even smaller. Once the yellow and orange curves are combined, the dip in that wavelength-band which parallels that in the hue-discrimination curve would disappear, making the range-curve flatter than it now appears. This change is just one of several which are likely to occur with changes in instructions in general or judgment categories in particular, and further experimentation is required to provide answers to some of these questions.

---

[19] Halsely and Chapanis, On the number of absolutely identifiable spectral hues, *J. opt. Soc. Am.*, 41, 1951, 1057-1058.

[20] E. A. Alluisi, Conditions affecting the amount of information in absolute judgments, *Psychol. Rev.*, 64, 1957, 97-103.

[21] Ekman, *op. cit.*, 1954, 467-474.

SUMMARY

By the method of absolute judgments and under two levels of luminance, data were obtained which divided the spectrum into six frequency-distributions, representing the judgmental categories of *violet, blue, green, yellow, orange,* and *red*. The *O*s were permitted to give qualifying responses after they had given the category response, and cumulative judgment-times were recorded for the latter. The two sets of six-category name-curves show little variation with luminance-level. The *PSE*s between any two adjacent name-curves occurred, in all cases, at the 50% point, and qualifying responses, when given, usually compounded adjacent category-names. Judgment-time and frequency of qualifying response were positively related; that is, exposures which gave long judgment-times also tended to evoke a high frequency of qualifying responses. Moreover, the spectral regions in which the frequency of occurrence of a categorical response was 100% tended to show short judgment-times and low frequencies of qualifying response, while these two measures tended to rise at the 50% points. Both auxiliary measures also pointed to spectral areas in which naming was problematical even when the frequency of the category-name was 100%.

A consideration of the yellow-orange dichotomy raised the question of whether it is not, in fact, a single category. A resemblance between color-naming and hue-discrimination was found to be suggestive, but not conclusive. The parallelism between the two functions might disappear under certain conditions of change in the instructions that would lead, for example, to a condition in which the yellow and orange curves became one.

THE INFLUENCE OF ASSUMED SIZE ON APPARENT DISTANCE

By WILLIAM EPSTEIN, University of Kansas

Every familiar object may have an 'assumed size.' This refers to "the entirely subjective sense of size which the observer might relate to a specifically characterized physiological stimulus-pattern."[1] It is stated frequently that, under certain conditions, the apparent distance of an identified object is related to the assumed size of the object. This relationship can be formulated as follows: "Changes in the assumed size of an object whose retinal size remains constant will result in appropriate changes in the apparent distance of that object."[2] This proposition seems to have a basis both in daily experience and in our thinking about the size-distance relationship.[3] Nevertheless, as noted elsewhere,[4] there is very little experimental evidence for the stated relationship between assumed size and apparent distance.

An experiment by Hastorf is cited often as a demonstration of this relationship.[5] Hastorf purported to show that under reduced conditions of viewing, the apparent distance of a single, simple geometric form will vary systematically with differences in the sizes which *E* suggests for the given stimulus. In view of the scarcity of experimental evidence bearing on the hypothesis of assumed size, it is worthwhile to reëxamine Hastorf's procedures. This examination results in the following observations:

(1) No effort was made to obtain a response which reflected *O*'s immediate perceptual experience. Instead, *O* was allowed a relatively extended interval during which he made a distance-judgment with which he was satisfied. Under these conditions, many extraneous factors may have affected the results.

---

* Received for publication May 18, 1961. This study was supported by research grants from the NIMH of the U. S. Public Health Service (M-4153), and the General Research Fund of the University of Kansas. Grateful acknowledgment is made to Raymond Engstrand who collected the data and performed the statistical computations.

[1] A. H. Hastorf, The influence of suggestion on the relation between stimulus size and perceived distance, *J. Psychol.*, 29, 1950, 195.

[2] William Epstein, John Park, and Albert Casey, The current status of the size-distance hypotheses, *Psychol. Bull.*, 58, 1961, 503.

[3] This proposition should be distinguished from the hypothesis which deals with discrete changes in the size of the retinal image of an object whose known size remains constant. See Epstein, Park, and Casey, *op. cit.*, 503; William Epstein, The known size-apparent distance hypothesis, this JOURNAL, 74, 1961, 333-346.

[4] Epstein, Park, and Casey, *op. cit.*, 503.

[5] Hastorf, *op. cit.*, 195-217.

(2) An indirect measure of apparent distance was used. The *O* was asked to vary the *size* of the stationary standard-object so that it appeared at the same distance as another point (pole) in the comparison-field. The validity of this procedure is limited by the reliability with which *E* can interpret the physical size-setting to represent *O*'s perception of distance. When variations in distance are under investigation, it is desirable that *O* be allowed to respond by varying the distance of a comparison-stimulus.

(3) No measure was obtained of the apparent size of the standard under the different conditions of 'suggested size'. In order to identify the variations in apparent distance as perceptual effects, it is necessary to know whether appropriate changes in apparent size accompanied the different suggested sizes.

(4) Only two points on the assumed dimension of size were sampled ('large' and 'small' meaning). A third point midway between the two extreme sizes would contribute considerably to our knowledge about the form of the size-distance relationship.

(5) Successive judgments were obtained from the *same O* under the two suggested-size conditions. This procedure admits the possibility that the differences in apparent distance may result from inter-size judgmental or memorial comparisons.

The objective of the present study was to investigate the relationship between assumed size and apparent distance in an experiment which takes into account the five observations given above.

Method: (1) Apparatus. The apparatus was a modified version of the Thereness-Thatness Table described in detail in Ittelson's manual.[6] This apparatus consisted of two adjacent alleys separated by a partition. The standard-alley contained the standard-stimulus only. The lighting of this alley could be varied from total darkness to normal illumination. In the former case the standard was illuminated by a beam of light which was restricted entirely to the surface of the standard. The comparison-alley contained a pack of cigarettes mounted on a track and was illuminated normally. The *O* could vary the radial distance of the pack continuously by turning a wheel which moved the pack along the track. This wheel made no sound and was not visible to *O*. In addition, a black screen could be mounted on the track at any desired distance. A specially devised projector and lens-system enabled *E* to project a circular disc of light on the screen. The size of the disk could be varied continuously by *O* without varying its intensity. This arrangement was used to obtain measures of the apparent size of the standard.

When *O* was in the viewing position he could see into both alleys simultaneously. An arrangement was contrived whereby the standard-alley could be viewed only monocularly through an artificial pupil, while the comparison-alley could be viewed binocularly at the same time. This arrangement could be modified easily to permit simultaneous, unimpeded binocular viewing of both alleys.

(2) Materials. The main materials were three circular standard-stimuli of equal diameters. One was a normally-sized (diameter = 2.4 cm.) photograph of a quarter *(Q)*. A second standard was a photograph of a dime *(D)* enlarged to the size

---

[6] W. H. Ittelson, *The Ames Demonstrations in Perception*, 1952, 21-24.

of a quarter. The third was a photograph of a half-dollar *(H)* reduced to the size of a quarter. Each of these photographs was pasted onto a disk cut from cardboard and then mounted on the top of wire-thin, black rods. Under the experimental conditions the standards appeared to *O*s as real coins.

(3) Observers. The *O*s were 45 paid volunteers from the undergraduate students who had no prior experience as experimental *O*s. None had any courses in psychology or acquaintance with the question under investigation. None of the *O*s wore corrective lenses.

(4) Procedure. The *O*s performed individually. Each *O* was assigned to one of three standards in rotation in the order of their appearance in the laboratory. The following procedure was used with *O*s assigned to the 'dime.' *O* was blindfolded, and led into the experimental room. After being seated and placed in the desired viewing position, *E* describe the general nature of the experiment. He was then shown the comparison-alley and performed various practice tasks with the comparison-stimuli. Then one ascending and one descending judgment of the apparent distance of the 'dime' were obtained under totally reduced conditions, *i.e.* the dime was presented in darkness and was viewed monocularly through an artificial pupil. The *O* made his judgments by moving the pack of cigarettes in the comparison-alley along the track until it appeared to him to be at the same distance as the dime. Then the dime was presented again, and *O* judged its apparent size. For this purpose the screen was raised in the comparison-alley and placed at *O*'s mean apparent distance-setting for the dime. *O* made one ascending and one descending judgment by varying the size of the comparison-disk of light until it appeared equal to the dime. Following this, *O* made apparent distance- and size-judgments of the dime under conditions of unimpeded, binocular viewing, *i.e.* all cues were restored.

The *O*s assigned to the 'quarter' *(Q)* and 'half-dollar' *(H)* proceeded in the same way except that these *O*s judged the 'quarter' or 'half-dollar.' In all cases the standard was located at a distance of 135 cm. from *O*'s eyes.

The instructions for the distance- and size-judgments were intended to induce a 'phenomenal attitude.' *O* was encouraged to base his setting on his first perceptual impression.

The experiment was concluded with an interview. The interview sought to clarify two aspects of the situation: (1) Did the settings reflect direct perceptual experience or were they secondary outcomes of deliberate judgmental efforts? (2) Did the apparent size-judgments represent the size of the coins as it appeared during the judgment of distance?

Results. Since there was little difference between ascending and descending judgments, the results of the two judgments were averaged for each *O*. The main results are presented in Tables I and II. The data in Table I show that with restricted monocular observation the half-dollar (*H*) was localized most distantly and the dime (*D*) appeared at the shortest distance. The apparent distance of the quarter (*Q*) was between that of the dime and the half-dollar. An analysis of variance yielded a significant *F* of 17.59 ($df = 2/42$, $P < 0.001$). Duncan's multiple-range test with $\alpha = 0.05$ and $df =$

42 was used to investigate the differences among the three means.[7] The shortest significant ranges were $R_2 = 17.86$ and $R_3 = 18.81$. Since $H - D$ $H - Q$ and $Q - D$ exceeded their respective ranges, we may conclude that these differences were significant. No statistical analyses were performed on the comparable data for unimpeded, binocular observation. An inspection of the data in Table I shows that the mean distance-judgments were almost identical for the three standards, and that they differed very little from the objective distance of 135 cm.

TABLE I

APPARENT DISTANCE OF STANDARD STIMULI WITH MONOCULAR AND BINOCULAR OBSERVATION

(Physical distance = 135 cm. All entries are in cm.)

Standard	Monocular observation		Binocular observation	
	Mean	*SD*	Mean	*SD*
Dime	112.74	23.9	135.70	2.57
Quarter	132.73	17.5	135.44	1.88
Half-dollar	164.19	29.7	135.40	1.92

TABLE II

APPARENT SIZE OF STANDARD STIMULI UNDER MONOCULAR AND BINOCULAR OBSERVATION

(Physical size = 2.4 cm. All entries are in cm.)

Standard	Monocular observation		Binocular observation	
	Mean	*SD*	Mean	*SD*
Dime	1.96	0.34	2.47	0.18
Quarter	2.55	0.33	2.52	0.04
Half-dollar	2.93	0.73	2.47	0.16

Table II contains the mean judgments of apparent size. The data show that with monocular, restricted observation the half-dollar was judged to be largest, while the dime appeared to be smallest. The apparent size of the quarter was between that of the dime and the half-dollar. An analysis of variance yielded a significant F of 12.78 ($df = 2/42$, $P < 0.001$). Duncan's test was used again ($\alpha = 0.05$, $df = 42$) to investigate the differences among the three means. The shortest significant ranges were $R_2 = 0.37$ and $R_3 = 0.39$. Since $H - D$, $H - Q$, and $Q - D$ exceeded their respective ranges, we may conclude that these differences were significant. No statistical tests were performed on the comparable data for binocular observation. An inspection of the data in Table II shows that the mean judgments of size were almost identical for the three standards, and that they differed very little from the physical diameter of 2.4 cm.

---

[7] A. L. Edwards, *Experimental Design in Psychological Research*, 1960, 136-140.

Tables I and II show that there were considerable differences in variability between the measures for monocular and binocular observation. For apparent distance the *SD*s are in a ratio of more than 10 to 1, and for apparent size the ratios vary from 8 to 2. These differences in variability are not unexpected. The relatively small *SD*s for apparent distance and size under conditions of unimpeded binocular vision reflect the fact that many cues for distance were available and, therefore, all *S*s gave judgments which were nearly veridical. The elimination of distance-cues under the condition of monocular observation created, however, a situation of indeterminate distance, *i.e.* apparent distance unspecified by the cues which are normally available. Within the range of possible apparent distances the localization of the standard is determined by the factor of assumed size. Since this factor depends on *remembered* size it will vary somewhat among *O*s (and perhaps for the same *O* on different occasions).[8] This variability in assumed size presumably can produce variability in apparent distance. The variability in apparent distance is not, however, in itself of importance. The significant consideration concerns the relationship between the variations in assumed size (apparent-size judgments) and variations in apparent distance.

Next, the relationship between the apparent distance and apparent size of each of the three standards under monocular observation was assessed. For this purpose the Spearman rank-correlation coefficient (corrected for ties) was used.[9] This analysis revealed the positive correlation between apparent size and apparent distance for each standard which was theoretically expected. *Rho* was 0.77 ($P < 0.01$), 0.56 ($P < 0.05$) and 0.76 ($P < 0.01$) for the dime, quarter, and half-dollar, respectively.

In Fig. 1 the mean judgments of apparent distance can be compared with the values which are required if apparent distance is, in fact, determined by the assumed size of the standard. In this experiment the apparent size-judgments can be taken to represent the assumed sizes of the standards. With this in mind, the requirements of the assumed-size hypothesis can be determined by the following simple formula:

$$ad = (as/ps)\ (pd)$$

---

[8] A. C. McKennell, Visual size and familiar size: Individual differences, *Brit. J. Psychol.*, 51, 1960, 27-35.

[9] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, 206-213.

where *ad* = theoretically required apparent distance; *as* = apparent (judged) size; *ps* = physical size; *pd* = physical distance.

Solving this formula for the appropriate values yielded the theoretical values of apparent distance of 110.70, 141.75 and 164.70 cm. for the dime, quarter, and half-dollar, respectively. These values have been plotted in Fig. 1 along with the values which were actually obtained. Inspection of these two curves reveals a striking degree of agreement. For the dime the difference between the obtained and theoretical values was 2.04 cm. (1.80%), while the discrepancy for the half-dollar was 0.51 cm. (0.03%). The deviation for the quarter was somewhat greater. The difference of 9.02 cm. (6.03%) represents a slight underestimation (2.27 cm.) of distance instead of the overestimation (6.75 cm.) which is required theoretically.

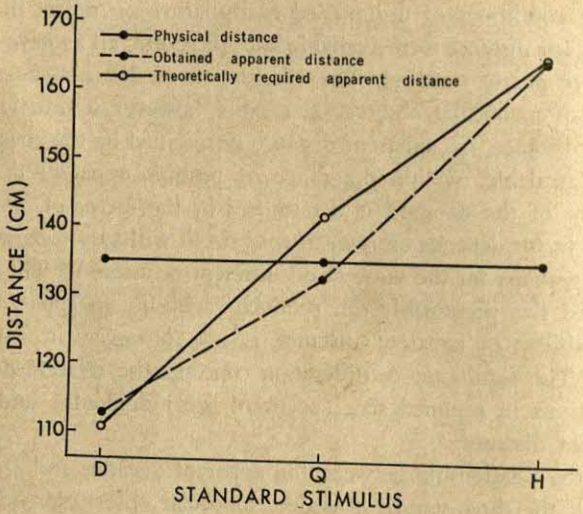


FIG. 1. JUDGMENTS OF APPARENT DISTANCE COMPARED WITH PHYSICAL DISTANCES AND THE THEORETICALLY PREDICTED VALUES
(For derivation of predicted values in text.)

One reservation must be imposed on the conclusion implied by the results of the preceding analysis. Although the analyses based on grouped data confirmed impressively the assumed-size hypothesis, equally strong confirmation was not present in the individual data. The discrepancies in the individual size-distance pairings may be illustrated by referring to the results of one *O* under Condition Q. This *O* ranked first in distance setting, *i.e.* he judged the quarter to be nearer than any other *O* in the group, but he ranked eleventh (largest) in size-setting. Applying the formula which was presented earlier to this *O*'s data revealed a deviation of 56.60 cm. (35.5%) from the theoretical requirements. Thus, for this *O*, and for all other *O*s also, the precise quantitative linkage between apparent size and apparent distance was not observed. This fact is reflected in the data in Table III, which gives the frequency, sign, and mean magnitude of the deviations from theoretical expectancy under conditions of restricted, monocular observation.

The following results of the postexperimental interview tended to confirm the validity of the experimental operations. All *O*s asserted that their size- and distance-settings were based on immediate perceptual experience. This assertion received independent support from the observation that the *O*s completed their judgments fairly rapidly. The intervals required for a single complete judging operation ranged from 8-14 sec. Most of this time was consumed by moving the distance-comparison or adjusting the size-comparison from its starting point. Spontaneous comments indicated that *O*s in this experiment found the instructions easy to follow. The judgments requested were considered 'natural.' Most *O*s indicated that they would not have tried to utilize any other method since introduction of other considerations would have complicated the matter for them. All *O*s felt that the same coin was being observed during the judgments of distance and size. Furthermore, with the exception of two *O*s, all *O*s felt the size-judgments could be taken to represent the apparent size of the standard when its distance was being judged. The two *O*s who indicated that this was not the case could not specify the nature of the difference.

TABLE III

DISTRIBUTION AND MAGNITUDE OF DEVIATIONS OF OBTAINED APPARENT DISTANCE FROM THEORETICAL DISTANCE

	Dime	Quarter	Half-Dollar
No. of positive deviations*	10	4	9
Mean magnitude of positive deviations (cm.)	11.81	15.52	13.00
SD of positive deviations	6.13	3.44	15.69
Av. positive deviations (%)	11.04	13.45	9.20
No. of negative deviations†	5	11	6
Mean magnitude of negative deviations (cm.)	16.31	20.22	21.04
SD of negative deviations	13.48	17.37	13.11
Av. negative deviations (%)	13.93	13.16	10.51

* A positive deviation refers to a distance-judgment which exceeded the theoretical requirement.

† A negative deviation refers to a distance-judgment which fell short of the theoretical requirement.

Discussion. The over-all picture presented by the preceding analysis represents an impressive confirmation of the hypothesis about the relationship between assumed size and apparent distance. The apparent distance of the standard was shown to vary systematically with the identified characteristics of the object. Furthermore, apparent size also was shown to vary with the identity of the standard.[10] The direct proportionality between apparent size and apparent distance which is indicated by the fairly high positive correlations between size and distance, and also by the curves in Fig. 1, provides compelling additional support for the hypothesis.

It should also be noted that these effects cannot be attributed to differences between the three standard-situations, whether these differences are presumed to be proximal or distal. This observation (if proven valid) is significant because several investigators have argued that the effect of known size on apparent distance may be reinterpreted as a relative size-effect.[11] In the present experiment the distal and proximal stimulus-situation remains constant for all three standards.

This consideration leads to the following question: Why has the assumed-size hypothesis received confirmation while most investigators have failed to support the known-size hypothesis? It is true that the respective hypotheses refer to different stimulus-situations. The assumed-size hypothesis applies to cases in which retinal size remains constant while assumed size varies. The known-size hypothesis refers to instances in which retinal size varies while known size remains constant. Nevertheless, both hypotheses are derivations from the general hypothesis of size-distance invariance, and, therefore, it is reasonable to expect that they would share a common fate.[12] A resolution of this apparent paradox must await further experimentation.

Another question that remains concerns the individual deviations from the theoretical requirements. It is possible that these deviations are due to the fact that size-judgment obtained some time *after* the distance-judgment can never correspond perfectly with the apparent size of the standard *during* the judgment of distance. While such assumed discrepancies might be expected to balance out in the grouped data, they would lead to errors in the evaluation of the individual data. This interpretation is only conjectural, and, for the present, the question of individual deviations from the theoretical requirement must remain unanswered.

---

[10] The tendency of apparent size to correspond with the real life size of the standards was, however, incomplete. The 'dime' was judged to be larger than a real dime, and the half-dollar was judged to be smaller than the actual size of the coin. Thus, the judgments converged slightly in the direction of the physical dimensions of the standard.

[11] C. B. Hochberg and J. E. Hochberg, Familiar size and the perception of depth, *J. Psychol.*, 34, 1952, 107-114; W. C. Gogel, B. O. Hartman, and G. S. Harker, The retinal size of a familiar object as a determiner of apparent distance, *Psychol. Monogr.*, 71, 1957, (No. 442), 1-16; Epstein, *op. cit.*, 333-346.

[12] For a more complete exposition of this point, see Epstein, Park, and Casey, *op. cit.*, 503.

SUMMARY

This experiment investigated the relationship between assumed size and apparent distance. The *O*s judged the apparent distance and size of mounted photographs of a dime, quarter, and half-dollar under reduced viewing conditions and unimpeded, binocular conditions. The photograph of the quarter was life-size. The 'dime' was enlarged to the size of the quarter while the 'half-dollar' was reduced to the size of the quarter. All the standards were located at the same physical distance from *O*. A different group of *O*s judged each standard. For the binocular conditions the *O*s' judgments of size and distance did not differ significantly from the objective size and distance of the standards. The results for the reduced conditions confirmed the hypothesis which relates assumed size to apparent distance. The apparent distance and apparent size of the coins covaried systematically in the manner predicted by the hypothesis.

Several important questions remain, however, to be answered concerning the bearing of this hypothesis on the general relationship between size and distance. In addition, a question was raised by the occurrence of individual deviations from the theoretical requirements.

TRANSFER FROM FAILED PAIRS AS A TEST OF ONE-TRIAL VS. INCREMENTAL LEARNING

By MARIAN SCHWARTZ, University of California, Berkeley

According to incremental theories of paired-associate learning, an association between stimulus and response is formed on the first learning trial, and each succeeding trial strengthens this association.

Opposing the traditional incremental view are the one-trial learning theories, which hold that the association between stimulus and response increases from zero to maximal strength on a single trial. Repetition plays no role in the formation of associations, other than that of providing the occasion for new ones to be formed, each on a single trial.[1] Repeated reinforcements simply give repeated opportunities for the formation of an association.[2]

If a correct response is given for the first time after *N* learning-trials (for any *N* greater than zero), the incremental interpretation would be that the cumulative effect of *N* trials raised associative strength to the level needed for performance; according to the hypothesis of one-trial learning, associative strength increased from zero to maximum on the *N*th trial. Failure after *N* trials indicates, from the incremental point of view, associative strength below the level required for performance but above zero; the one-trial interpretation would be that the association between stimulus and response was still at zero strength, *i.e.* that no association had been formed.

This experiment measured the associative strength of pairs failed after one learning trial by testing for transfer from these pairs to later learning. For an Experimental Group, the response-members of failed pairs were interchanged. For a Control Group, these pairs remained unchanged. Since any associative strength which may have accrued to the failed pairs on the initial trial should hinder learning of the changed items, and help learning of the unchanged items, incremental theory predicts that the Control Group's performance on the unchanged failed pairs should surpass the Ex-

---

* Received for publication August 8, 1961. The writer is indebted to Dr. Leo Postman for his suggestions and criticisms.

[1] Irvin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193.

[2] W. K. Estes, B. L. Hopkins, and E. J. Crothers, All-or-none and conservation effects in the learning and retention of paired associates, *J. exp. Psychol.*, 60, 1960, 329-339.

perimental Group's performance on the corresponding changed pairs; one-trial learning theory predicts no difference in performance between the two groups.

Method: (1) Design. All *Ss* had one study-trial on a list of paired associates, followed by a test-trial. For the Experimental Group, response-members were interchanged for all pairs failed on the test-trial. For the Control Group, the failed pairs were left unchanged. Each *S* then was given one more study-trial, followed by two test-trials.

The Experimental and Control Groups were treated alike with respect to the pairs given correctly on the first test: if two, three, or four pairs were correct, two had their response-members interchanged; if five were correct, three were changed. (No *S* had more than five correct responses on the first test.)

The responses of correct items were interchanged only with those of other correct items; and the responses of items failed only with those of other failed ones. Interchanging was done according to a different random arrangement for each *S*.

While the items were being re-paired, the *S* was occupied with a 5-min. task which required him to sort pictures into five piles on an esthetic rating scale.

(2) Materials. A list of 18 number-word pairs was used. The stimulus-members were the numbers from 51 through 68. The response-members were high-frequency, three-letter nouns, classified AA or A in the Thorndike-Lorge general word-count.[3] No two words had the same initial letter.

To permit later interchange of the response-members, the numbers and words were typed on separate 3 × 5-in. cards. The number was placed slightly to the left of center on the stimulus-card. The word was typed at the left-hand edge of the response-card, which was folded into 3 × 2½-in. halves and taped along the top edge. The right-hand side of the stimulus-card was then sandwiched between the folded halves of the response-card, aligning the number and word. On the test-trials, blank cards were substituted for the response-cards.

(3) Procedure. The cards were shown through a window in a screen separating the *S* and the *E*. On all trials, the rate of presentation was 6 sec., with no interval between pairs. The intertrial-intervals were about 10 sec. each, except for the period between the first test and the second study-trial, which averaged 6 min. for both groups.

Instructions were the same throughout for the Experimental and Control Groups. *S* was advised that the stimuli would be numbers between 51 and 68, the responses would be three-letter words, and the task would be to remember the word that went with each number. To discourage the strategy of concentrating on some pairs and ignoring others, the *S* was instructed always to concentrate on the pair being shown and not to spend time reviewing other pairs. Before starting the esthetic rating task, the *S* was told that trials would later be resumed with somewhat different pairs. The instructions for the second study-trial stated that the same numbers and words as before would be shown, that in some (but not all) cases a word which formerly went with one number would now be shown with another number, and that the task would be to remember the word that went with each number in this second series. These instructions were meant to engender in both groups the expectation that some pairs would be changed and some would not, an expectation which was supported whenever possible by the interchanging of responses among some of the pairs correct on the first test. On none of the test-trials were the *S*s told whether their responses were right or wrong.

---

[3] E. L. Thorndike and Irving Lorge, *A Teachers' Word Book of 30,000 Words*, 1944.

Five pairings of the stimulus- and response-units were used on the first study-trial. The pairings were random except that duplications were avoided. Pairs were presented in a different order on each of the study-trials and test-trials. The orders of presentation were random, subject to the following restrictions: (a) no stimuli were shown consecutively if they were naturally consecutive numbers (*e.g.* 61, 62), if they shared the same second digit, or if they were shown consecutively in another order; (b) no stimulus was used in a terminal position in more than one order; and (c) at least three other pairs intervened between presentation of a given pair on one trial and on the next trial.

(4) Subjects. The *S*s were 50 college students from undergraduate courses in psychology; 25 assigned to the Experimental and 25 to the Control Group. Assignments to the two groups and to the five initial pairings were randomly distributed over each block of 10 *S*s, and assignments to the orders of presentation to be used on each trial were randomized over the two groups, so that each pairing and each starting order were used for five Experimental and five Control *S*s.

Results. The number of correct responses on Test 1 ranged from 0 to 5 for the Control, and from 0 to 4 for the Experimental Group. The mean correct was 2.32 for the Control and 1.80 for the Experimental. The difference between the two groups was not significant ($t = 1.44$, $P > 0.10$).

The mean number of critical pairs correct on Tests 2 and 3 is shown in Table I. The critical pairs are those failed on Test 1 and then changed (for the Experimental Group) or left unchanged (for the Control Group). After a Freeman-Tukey square-root transformation of the scores to correct for heterogeneity of variance, the mean number of critical pairs correct on Test 2 was found to be significantly higher for the Control than for the Experimental Group ($t = 2.06$, $P < 0.05$).

TABLE I

MEAN NUMBER OF CRITICAL PAIRS CORRECT ON TESTS 2 AND 3

Test	Group	Mean	*SD*	Range
2	Control	3.88	2.41	0–8
	Experimental	2.52	1.68	0–6
3	Control	4.04	2.05	1–8
	Experimental	2.44	1.60	0–6

Since the difference between the two groups on Test 1, although not significant, was in favor of the Controls, correlations were computed between the scores on Test 1 and Test 2 to see whether it would be advisable to adjust the latter scores in an analysis of covariance. Since the correlations were small and not significant ($r = 0.13$ for the Control and 0.06 for the Experimental Group), no adjustment was deemed necessary.

Difficulty of items. While the superior performance of the Control Group on the critical pairs is consistent with the incremental prediction, another explanation must be considered: the critical pairs for the Experimental Group might have been, aside from any transfer effects, inherently more difficult than those for the Controls. A priori, the reverse seems more likely. Assuming differences in difficulty among the original pairs, those correct on Test 1 should have been relatively easy and those failed, relatively hard. Randomly interchanging the response-members of these relatively hard pairs should have resulted in regression toward an average level of difficulty. Thus, the Experimental Group's changed pairs should have been easier rather than harder than the unchanged pairs of the Controls (again, aside from any transfer-effects).

This reasoning was borne out by an analysis of the critical items. A rough measure of the difficulty of the original pairs was derived by counting the total number of *Ss* who responded correctly to each item on Test 1. (The Experimental and Control Groups were combined for this count.) Pairs given correctly by no *S* or one *S* were labeled 'hard', those given correctly by two or more *Ss*, 'easy.' By this measure, 67% of all pairs shown on the first study-trial were 'hard.' This percentage applies to both groups since the original pairs were the same for both. Considering only the pairs failed on Test 1 (*i.e.* the Control Group's critical pairs, and the corresponding pairs for the Experimental Group before interchanging), the percentage of 'hard' pairs turned out to be 72% for both groups. The new combinations produced by the re-pairing procedure could be classified by this measure of difficulty only when the changed pair happened to be the same as an original pair in another pairing. This was the case for 113 of the Experimental Group's 405 critical pairs. Of these 113 pairs, 58% were 'hard.' It is clear that, at least according to the rough measure used, the critical pairs of the Experimental Group were not more difficult than those of the Controls, and that the superior performance of the Control *Ss* occurred in spite of, rather than because of, any differences in difficulty of items.

Conditional response-probabilities. Another approach to the problem of testing one-trial vs. incremental predictions is by way of conditional response-probabilities.[4] In the present experiment, the relevant data are the critical pairs in the category C_3: N_2 (correct on Test 3, given that the item was not correct on Test 2) as a proportion of all critical pairs not correct on Test 2; and the critical pairs in the category C_3: C_2 (correct on Test 3, given that the item was correct on Test 2) as a proportion of all critical pairs correct on Test 2. It may be predicted on the incremental hypothesis that both these proportions should be higher for the Control Group, which had two study-trials on the critical items, than for the Experimental Group, which had only one. Implicit in this prediction are the assumptions that there is a threshold of response evocation, that items oscillate back and forth across this threshold from moment to moment, and that the range of oscillation is a function of the associative strength of the item. Therefore, a critical item which was failed on Test 2 should be closer to threshold, and thus more likely to oscillate above threshold on Test 3, in the control than in the experimental condition. Similarly, in the control condition an item correct on Test 2 should be more likely to remain above threshold on Test 3, *i.e.* retention of learned items should be greater after two study-trials than after one.

---

[4] W. K. Estes, Learning theory and the new "mental chemistry," *Psychol. Rev.*, 67, 1960, 207-223; Estes, Hopkins, and Crothers, *op cit.*, 329-339; Leo Postman, One-trial learning, in C. N. Cofer and B. S. Musgrave (eds.), *Problems and Processes in Verbal Behavior and Learning*, in press.

On the one-trial learning hypothesis that failed items have zero and correct items maximal associative strength, regardless of the number of prior learning trials, there should be no difference between the Experimental and Control Group in either of these response probabilities.

As shown in Table II, the number of C_3: N_2 responses as a proportion of all N_2 responses is 0.081 for the Controls and 0.053 for the Experimental Group.[5] (Calculated as a proportion of all overt responses made on Test 3 to N_2 items, C_3: N_2 responses are 0.300 for the Controls and 0.189 for the Experimental Group.) The number of C_3: C_2 responses as a proportion of all C_2 responses is 0.794 for the Controls, and 0.683 for the Experimental Group. All differences are in the direction predicted by the incremental hypothesis.

Joint probabilities. Another test of incremental theory has been suggested for the situation in which a reinforced trial is followed by two consecutive unreinforced test-trials.[6] According to Estes, incremental theory (specifically, a simple linear model) requires that two combinations of responses be equally probable: (a) correct on the first of the two consecutive tests, not correct on the second; and (b) not correct on the first, correct on the second. In the present study, the relevant comparison is between the proportions of all critical pairs falling in the categories N_3: C_2 and C_3: N_2. As shown in Table II, these proportions differ by only .01 for the Controls and by less than that for the Experimental Group.

---

[5] A breakdown by difficulty of items shows that the Control Group's C_3:N_2 proportions were 0.141 for 'easy' and 0.062 for 'hard' items. This finding is consistent with the assumptions of incremental learning and a response threshold: after a given number of reinforcements, 'easy' items should be oscillating closer to threshold than 'hard' items, so should be more likely to exceed threshold on the next test. No similar breakdown was feasible for the Experimental Group due to the small number of C_3:N_2 items that could be evaluated for difficulty.

[6] Estes, New developments in statistical behavior theory: differential tests of axioms for associative learning, *Psychometrika,* 26, 1961, 73-84.

TABLE II

NUMBER AND PROPORTIONS OF CRITICAL PAIRS WITH VARIOUS COMBINATIONS OF RESPONSES ON TESTS 2 AND 3

Combination of responses	Number		Proportion					
	cont.	exp.	of all C_2*		of all N_2*		of all pairs†	
			cont.	exp.	cont.	exp.	cont.	exp.
C_3: C_2	77	43	.794	.683	—	—	.196	.106
N_3: C_2	20	20	.206	.317	—	—	.051	.049
C_3: N_2	24	18	—	—	.081	.053	.061	.044
N_3: N_2	271	324	—	—	.919	.947	.691	.800

* Conditional response probabilities.
† Joint probabilities.

Intralist errors. Intralist errors (responses which could have been correct if given to another stimulus) are reported as percentages of opportunities for such errors, *i.e.* as percentages of all items not correct.

For the Experimental Group, the percentages of intralist errors on Tests 1, 2, and 3 were 19.4, 26.9, and 20.9, respectively. The corresponding percentages for the Controls were 18.1, 18.1, and 18.2. (On Tests 2 and 3, both errors and opportunities pertain to critical pairs only.) It is apparent that the re-pairing procedure for the Experimental Group did not depress production of intralist errors; therefore, the relatively low rate of correct responses after re-pairing cannot be attributed to a general decline in rate of response.

The fact that intralist errors are higher for the Experimental than for the Control Group on Tests 2 and 3 is consistent with an assumption of negative transfer in the experimental condition. The differences between the two groups, however, are not significant (for Test 2, after arcsine-transformation of the percentages, $t = 1.73$, $0.05 < P < 0.10$).

One type of intralist error is of particular interest. Experimental *S*s occasionally gave the original response, *i.e.* the response which would have been correct on Test 1, to a critical item. This occurred nine times on Test

2 and five times on Test 3, which is about as often as would be expected by chance.[7]

Thus, an analysis of intralist errors provides no evidence for specific negative transfer from the first to the second study-trial for the Experimental Group; the comparison between the two groups in number of errors made is inconclusive.

Discussion. The superior performance of the Control Group on the critical pairs, in spite of the fact that differences in difficulty of items favored the Experimental Group, argues that some associative strength accrued to the pairs failed on Test 1 and affected the course of later learning. How much of the difference between the two groups may have been due to negative transfer in the experimental condition, and how much to positive transfer in the control condition, cannot be determined. Any transfer from the failed items is evidence, however, of associative strength greater than zero, and is therefore consistent with the incremental hypothesis and contrary to one-trial learning theory.

The supplementary evidence provided by conditional response-probabilities and joint probabilities supports the more direct comparison between the two groups in the number of correct responses on Test 2.

The design of the present study may be compared with one of Rock and Heimer, in which a list of eight pairs of nonsense-syllables was learned by the Experimental Group, with the response-members of failed pairs being interchanged after each trial.[8] No significant difference in trials (or errors) to criterion was found between the Experimental Group and a Control Group which learned the original list to criterion. The discrepancy between these results and those of the present experiment may reflect greater opportunity for selection of easy combinations by the Experimental *S*s in the study of Rock and Heimer.[9] It is likely that, on each trial, relatively easy combinations were learned and relatively hard ones were eliminated. Over a series of trials, the advantage to the Experimental Group from selection of easy items may have canceled the advantage to the Control Group from transfer.

---

[7] A chance-level for this type of occurrence was estimated by randomly assigning fictitious original response-members to the Experimental Group's critical pairs, and counting the number of times an intralist error matched the fictitious original response-member. For each pair, a different fictitious response was selected from a pool consisting of all original response-members of critical pairs for that *S* (excluding the actual second response-member for that pair, since it could not have been the original response-member), and was used for both Tests 2 and 3. The entire procedure was repeated three times, yielding three independent counts.

[8] Irvin Rock and Walter Heimer, Further evidence of one-trial associative learning, this JOURNAL, 72, 1959, 1-16.

[9] B. J. Underwood, Robert Rehula, and Geoffrey Keppel, Item-selection in paired-associate learning, this JOURNAL, 75, 1962, 353-371.

SUMMARY

Two groups were given a study-trial and a test-trial on a list of 18 number-word pairs. In the experimental condition, response-members of all pairs failed on the test were randomly interchanged. In the control condition, failed pairs were left unchanged. A second study-trial and two more test-trials followed.

Control *S*s surpassed Experimental *S*s in the number of critical pairs correct on the last two tests, in spite of the fact that differences in difficulty of items favored the Experimental Group. The proportions of responses correct on Test 3 were higher for the Controls both after correct and after incorrect responses on Test 2.

It was concluded that pairs failed after one study-trial had acquired some associative strength, which affected later performance through positive transfer in the control condition, or negative transfer in the experimental condition, or both. The findings are contrary to the hypothesis of one-trial learning that failed items have no associative strength, and support the opposed incremental view.

CHOICE REACTION-TIME AS A FUNCTION OF MEANINGFUL SIMILARITY

By NORMAN J. SLAMECKA, University of Vermont

The time required to make a choice between alternative stimuli has been shown to be directly related to the degree of physical similarity between the stimuli. Several studies exemplifying this relationship have used displays varying along such physical continua as wave-length, frequency, and linear extent.[1] All these types of stimuli allow for an independent, objective measurement of the distances between the alternatives. Non-sensory or psychological continua have not, however, been given the close attention that they merit, and we stress particularly the attribute of similarity of word-meaning.

Considering the salient role of words as discriminative stimuli governing human behavior, it is surprising that there is but one published report of the effects of the similarity of meaningful words upon latencies of responce.[2] In Dember's study, *S* was required to choose between two pairs of words, the one pair whose members appeared closer together in meaning. A pool of five words was used, from which pairs were formed, (viz. always, often, rarely, seldom, and never), and data were obtained from a single *S*. A significant negative relationship was found between reaction-time and perceived distance between the word-pairs. In spite of the elegance of demonstrating the relationship with an *N* of one, it seemed to us desirable to conduct a more extensive investigation of this problem, using both more *S*s and a wider range of stimulus-words, in order to extend the generality of the findings. Also, rather than asking *S* to choose between two pairs of words (a requirement in the Dember study, dictated by the use of the "Method of Cartwheels"), we decided to require a choice between two single words. Finally, since the concept of meaningful similarity is certainly not an unambiguous one, it seemed appropriate to break it down into the sub-aspects of denotative similarity (definition) and connotative similarity (properties and associations).

In each of the three experiments to be described, it was hypothesized that choice reaction-time for a pair of words would be positively related

---

* Received for publication October 25, 1960. This research was supported by the National Science Foundation.

[1] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology,* 1954, 33.

[2] W. N. Dember, The relation of decision-time to stimulus similarity, *J. exp. Psychol.,* 53, 1957, 68-72.

to the degree of meaningful similarity obtaining between those words. The problem of putting the words in acceptable categories in terms of their meaningful similarity was approached by referring the classification to each *S*'s own perception of the meanings of the words. In Experiments I and II, independent sortings for similarity were made by every *S* after his particular reaction-times were recorded. Thus, *S* himself determined which of three levels of closeness in meaning each word-pair represented, and it was against these individual sortings that the respective reaction-times were evaluated. In Experiment III, the criterion of meaningful similarity was the Osgood semantic differential scale-value of each word in a pair. These scale-values were obtained from the normative tables published by Jenkins, Russell, and Suci.[3]

EXPERIMENT I

Subjects. Twelve students from introductory psychology classes participated as *S*s as part of the course requirements. Each *S* took part in only one session, which included all conditions, thus giving a total of 12 replications.

Materials. A set of 35 3 × 5-in cards was prepared, on each of which were typewritten 3 adjectives in capital letters. The adjectives were placed in a triangular position, such that one word was centered above the lower two. With the aid of a dictionary, triads were chosen such that the lower two adjectives were either synonyms, or related in meaning to the upper word but not synonyms, or antonyms with respect to each other, and such that at least one of them was a synonym for the upper word. There were 12 cards for each relational category, and no adjective appeared more than once in the entire set of 36 cards. Regardless of this initial structuring, each *S*'s own subsequent sorting of the cards into High, Intermediate, or Low similarity served as the basis for the statistical analyses to follow.

Nine of the cards used, three from each level, were the following (the first word was in the upper position on the card, and the two hyphenated words appeared below):

High	Intermediate	Low
Unselfish: Generous-Giving	Icy: Frozen-Cool	Clear: Distinct-Obscure
Glorified: Praised-Honored	Brutal: Savage-Rude	Melancholy: Depressed-Joyful
Confused: Bewildered-Perplexed	Angry: Annoyed-Infuriated	Infinite: Limited-Endless

S controlled two telegraph keys, one for each hand, the lifting of which served to indicate his choice (left or right) of word for each trial, and at the same time broke the circuit with a Standard Electric timer, reading to 0.01 sec., thus registering his choice reaction-time.

Procedure. The set of cards was shuffled thoroughly at the start of each session.

---

[3] J. J. Jenkins, W. A. Russell, and G. J. Suci, An atlas of semantic profiles for 360 words, this JOURNAL, 71, 1958, 688-699.

S was informed that three different adjectives appeared on each card and that his task was always to choose that one of the two lower words which came closer in meaning to the upper word. Meaning was specified as the preferred dictionary definition. To eliminate false starts, *S* was requested, for the start of each trial, to depress both keys and to say "ready" when he was set to see the card. He was told to read all three words, make his best choice, and then to indicate it immediately by raising the appropriate hand from its key. Approximately one second after *S* said "ready," a cover in front of the card was dropped, revealing the card and at the same time starting the timer. Following every choice, the reaction-time was recorded. After the 36 reaction-times were obtained, all of the cards were given to *S* and he was asked to sort them into three equal piles on the basis of his perception of the degree of similarity (High, Intermediate, or Low) between the lower word-pairs of each card. *S* was allowed unlimited time for this task, and was also allowed to change cards from pile to pile at his discretion.

EXPERIMENT II

Another test of the hypothesis that decision-times are longer when the discrimination between two outcomes is more difficult was conducted. This experiment investigated reaction-latencies to the connotational dimension of meaning, exemplified as the properties or characteristics possessed by the referents of various nouns. For example, "brittle" is not the definition of "glass," but it is certainly a part of its connotational meaning, in the sense of referring to one of the common properties of glass.

Subjects. Another group of 12 students in psychology served for a single session, producing 12 replications of the experiment.

Materials. A set of 36 cards was prepared, each bearing three typewritten words placed in a triangular formation, as before. The uppermost word was an adjective referring to some physical attribute, and the lower two words were nouns that seemed to approximate any one of three levels of closeness to each other, with respect to degree of possession of the property in question. This preliminary selection of word-pairs was intended only to provide a reasonably wide range of stimuli at the outset, so that the ultimate sorting, which was the task of each *S*, might have adequate representation at each level. Twelve adjectives were used, each one appearing on three cards, associated with three levels of presumed word-pair similarity. No noun was used more than once. Nine of the cards actually presented, three from each level, were the following:

High	Intermediate	Low
Black: Soot-Coal	Black: Basement-Cave	Black: Cream-Tar
Smooth: Satin-Silk	Smooth: Velvet-Lawn	Smooth: Gravel-Paper
Thin: String-Wire	Thin: Tissue-Dime	Thin: Cow-Thread

Procedure. The rationale and general procedure followed in presenting the cards and in determining choice reaction-times were the same as those in Experiment I. *S* was instructed to choose that one of the two nouns on each card which more closely approximated or deserved the property expressed by the upper adjective. *S* indicated his choice by lifting his hand from the appropriate key. After the reaction-times were obtained, *S* was asked to sort all 36 cards into three equal piles according to his judgment of the degree of similarity between the nouns on each card. Thus, if both nouns were very similar to each other with regard to the attribute in question, that card would be put in the High category; if they were extremely dissimilar to each other on that attribute, the card would be classed as Low; and if both nouns exemplified the characteristic, but to moderately unequal extents, the card would fall in the Intermediate pile. *S* sorted each of the twelve triads of cards in turn (a triad was composed of cards having a common adjective) and was allowed to revise previous placings if he desired.

EXPERIMENT III

Experiment III utilized that attribute of connotational meaning measured by the Osgood semantic differential technique.[4] This aspect of meaning is not identical to the property-implications type examined in Experiment II, but refers essentially to the associative relations of a word to certain polar-opposite pairs of standard adjectives. Any given word can be assigned a number on a 7-point scale, with regard to its position between these polar opposites. Thus, this study is a test of response-latencies to word-pairs having known semantic-differential ratings.

Subjects. Another group of 12 naïve *Ss* was obtained from introductory classes in psychology, and each *S* participated individually in one session.

Materials. A set of 36 cards was again prepared, each bearing 3 words placed in the usual triangular arrangements. All of these words were taken from the normative tables of semantic profiles for 360 words.[5] The uppermost word on each card was one of the polar-opposites which form the terminal points of a semantic differential scale. For instance, from the polar pair "calm-excitable," the word "calm" was used. A total of 12 such upper adjectives appeared on the 36 cards, forming 12 triads, each member of which represented one of the three relative levels of lower word-pair similarity. In each case, the two lower words were those for which mean semantic scale-values with respect to the upper word, were available in the tables. High, Intermediate, or Low similarity-levels were established for these cards on the basis of the mean differences in semantic scale-values between the two lower words. There were 12 cards for each relative similarity-level. For the Low set, the mean semantic scale-value difference between the choice-words was 520.8, for the Intermediate set, it was 257.9, and for the High, it was 2.7. Since the similarity-levels were established beforehand and remained fixed in this experiment, it was possible to make certain that the choice-words comprising each level were comparable in Thorndike-Lorge frequency, lest the results be confounded by the factor of differential prior exposure.[6] The inclusion of words was therefore subject to this additional constraint. For the High level, Thorndike-Lorge frequencies were: AA = 9, A = 3, mean of all others, = 28.3. The Intermediate level had: AA = 9, A = 3, mean of all others = 22.3. The Low level had: AA = 9, A = 3, mean of all others = 23.9.

---

[4] C. E. Osgood, The nature and measurement of meaning, *Psychol. Bull.*, 49, 1952, 197-237.

[5] Jenkins, Russell, and Succi, *op. cit.*, 688.

[6] E. L. Thorndike and Irving Lorge, *The Teacher's Word-Book of 30,000 Words*, 1944. The designation "AA" refers to words occurring 100 or more times per million, and "A" is a count between 50 and 100 per million.

Nine of the cards, three from each level were the following:

High	Intermediate	Low
Wise: Flaming-Fleecy	Wise: Hasty-Music	Wise: Scientist-Steal
Beautiful: Sudden-Long	Beautiful: White-Tense	Beautiful: Starving-Flowers
Weak: Fragrant-Charming	Weak: Dawn-Late	Weak: Sickness-Trees

Inspection of these examples, particularly of the High and Intermediate levels, gives some indication of the difficulty of formulating verbally the attribute or attributes measured by the semantic differential technique.

Procedure. The same technique was used for presenting the stimuli and obtaining reaction-times as in Experiments I and II. The directions to *S* were simply to make a choice between the two words in favor of the word that seemed to be closer in meaning to the upper word, and no further elaboration of the notion of 'meaning' in the case was given. Contrary to what might have been expected, the *S*s did not question or appear to be puzzled about the task. After the reaction-times were obtained, the procedure was terminated. There was no sorting task for *S*, since the relative similarity levels were already determined by the selection of words from the norm tables.

RESULTS AND CONCLUSIONS

Each *S* provided 12 reaction-times at each level of similarity, self-determined in Experiments I and II and derived from the Semantic Differential in Experiment III. The median time (in sec.) for each *S* at each level of similarity was obtained, and the means of these values for all the *S*s in each of the experiments are found in Table I.

TABLE I

MEANS AND *SD*s OF MEDIAN REACTION-TIMES (IN SEC.) PER *S* AT EACH LEVEL OF SIMILARITY FOR EACH EXPERIMENT

	Similarity level					
Experiment	High		Intermediate		Low	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
I	2.86	2.57	1.87	0.65	1.53	0.40
II	2.30	1.12	1.70	0.47	1.35	0.28
III	2.17	0.69	1.60	0.38	1.22	0.24

Since inspection of the data showed that the means and *SD*s were correlated (as is usually the case with reaction-latencies), they were first subjected to a log transformation. Thereafter, a three-part analysis of variance was carried out for each experiment.[7] For all three analyses, the effect of levels of similarity, and also of *S*s, was significant beyond the 0.001

---

[7] A. E. Edwards, *Statistical Methods for the Behavioral Sciences,* 1954, 349.

level. For similarity-levels: Experiment I, $F = 14.0$; Experiment II, $F = 24.9$; Experiment III, $F = 66.3$ ($df = 22$ in each case). Tukey's gap test showed that the difference between adjacent means in every case was significant at least beyond the 0.05 level. For *S*s: Experiment I, $F = 10.0$; Experiment II, $F = 10.0$; Experiment III, $F = 13.2$ ($df = 11$ in each case). We therefore conclude that the central hypothesis under scrutiny in this series of experiments is strongly supported; that the phenomenon of a positive relationship between choice reaction-time and degree of stimulus-similarity applies to the dimension of meaningful similarity, and can be generalized both to denotative and connotative aspects of words.

When the median reaction-times for every individual *S* were examined, it was expected that there would be a fair number of reversals or disruptions in the order of magnitude from High to Intermediate to Low. The total number of reversals in Experiment I was, however, only three; in Experiment II, it was two, and in Experiment III, there was none at all. This fact attests to the stability of obtained relationship, both within as well as across *S*s.

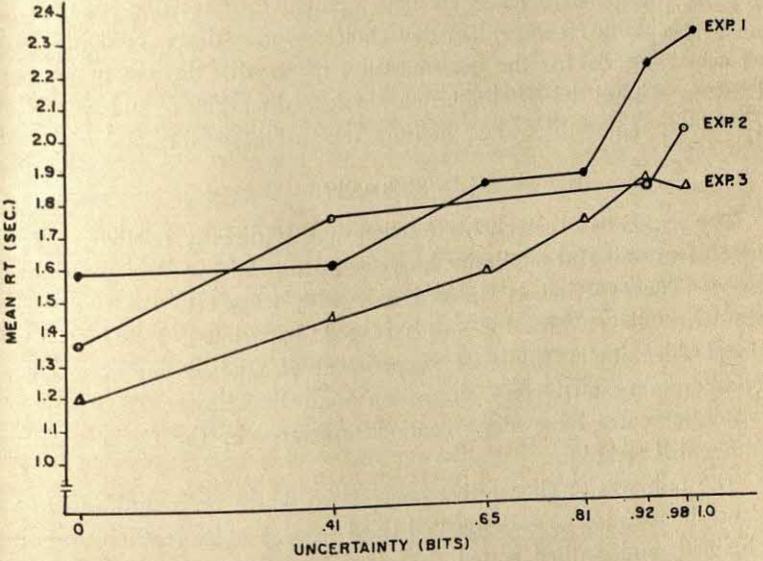


FIG. 1. RELATION BETWEEN UNCERTAINTY OF CHOICE AND MEAN CHOICE REACTION-TIME

The proportion of choices for one alternative or the other can also be examined as an objective measure of whether the choice is between widely different alternative or close ones. In each of the three experiments, a total of 12 *S*s made a choice for each card. Tabulation of the actual word chosen would result in seven proportions, ranging from 12-0 (or 0-12) to 6-6. In terms of an information measure, the uncertainty of choice for a 12-0 card would be zero bits, and it would increase through six steps to a maximal uncertainty of one bit for a 6-6 card. All seven choice-proportions were converted to such measures of uncertainty. The mean reaction-times for all the cards of a given uncertainty-level were calculated, and the relationship between these two variables was plotted for each experiment, as seen in Fig. 1.

To increase reliability, points were not plotted for those levels of uncertainty which contained less than two entries. Inspection of Fig. 1 reveals a consistent rise in the reaction-times with increasing uncertainty of choice within each of the three experiments. Thus, uncertainty measures and judged similarity-levels are both predictive of reaction-latencies.

In concluding, it is evident that choice reaction-times could be utilized as actual criteria for the determination of relative degrees of similarity between stimuli lying along non-physical dimensions, and with some confidence, since these data indicate reasonable stability and objectivity.

SUMMARY

The hypothesis that choice reaction-time is directly related to the degree of meaningful similarity between words was tested in three experiments. The experiments differed only with respect to the continum of similarity along which the words ranged. Experiment I studied denotational (definition) similarity; Experiment II studied one connotational (properties or attributes) dimension; and Experiment III investigated word-differences lying along continua measured by the Osgood semantic differential technique.

The task was to choose that one of two words on a card which came closer in meaning to a designated third word. Choice reaction-times were obtained, and *S* then sorted the cards into three categories expressing his perception of the degree of similarity between the two choice-words on each card. For Experiment III, the similarity-levels were predetermined from normative tables.

The results showed consistently longer reaction-times for the choices between more similar alternatives. Moreover, the proportion of choices for each card, expressed in terms of uncertainty, revealed a positive relationship between mean reaction-time and degree of uncertainty. It is suggested that choice reaction-times could be used as criteria for the determination of degrees of similarity between stimuli lying along non-physical dimensions.

TRANSFER, SYNONYMITY, AND ANTICIPATORY INTERVAL IN PAIRED-ASSOCIATE VERBAL LEARNING

By WILLARD N. RUNQUIST and MARY ANN MARSHALL, Pomona College

It has been generally accepted that in the learning of successive paired-associate lists, positive transfer will result when the stimuli of the two lists are identical and the responses are synonyms.[1] The evidence for this generalization, however, is more logical than empirical. It has been demonstrated that relative to the first list positive transfer increases as degree of synonymity increases.[2] These studies, however, have not controlled for non-specific transfer, *i.e.* learning to learn and warm-up; hence, it is not possible to specify the direction of transfer due to synonymity. In studies where such a comparison is possible, no positive transfer has been demontrated when responses are synonyms.[3] These studies, however, were not so designed that a direct comparison of the control-condition with the synonym-condition could be made. The purpose of the present studies was to determine the direction and amount of transfer when the stimulus-syllables are identical and the responses are synonyms.

In addition, the interval of anticipation was varied. Theoretical and empirical considerations lead one to suspect that this would be an important variable. If transfer in this situation is attributed to some kind of mediating association, it might be suspected that allowing a longer time-interval in which to anticipate the correct response would give a greater opportunity for the mediating response to occur. Empirically, Barnes and Underwood have shown that transfer to a second list of synonyms is essentially perfect when a written unpaced test of recall is used.[4]

Method. Sixty high school seniors were assigned randomly to one of three groups which varied in the anticipatory interval used in learning the second of two lists

---

* Received for publication June 7, 1961. This research was partially supported by a grant from the Claremont Graduate School.

[1] B. J. Underwood, *Experimental Psychology: An Introduction*, 1949, 303.

[2] C. E. Osgood, Meaningful similarity and interference in learning, *J. exp. Psychol.*, 36, 1946, 294-301; R. K. Young, Retroactive and proactive effects under varying conditions of response similarity, *ibid.*, 50, 1955, 113-118.

[3] B. R. Bugelski, and T. C. Cadwallader, A reappraisal of the transfer and retroaction surface, *ibid.*, 52, 1956, 360-367; B. J. Staner, Transfer in verbal learning as a function of stimulus and response similarity. Unpublished M.A. thesis, Northwestern University, 1956.

[4] J. M. Barnes and B. J. Underwood, Fate of first list associations in transfer theory, *ibid.*, 58, 1959, 97-105.

of eight nonsense syllable-adjective pairs.[5] These time-intervals were: 1, 2, and 4 sec. The first list was learned with a 2-sec. anticipatory interval for all groups. The second list consisted of four nonsense-syllables from the first list for which the responses were synonyms of the responses to the first list. The other four pairs were new syllable-adjective pairs. These latter items served as a control for non-specific transfer-effects. Items were counterbalanced by use of the same second list for all *Ss*, but two different first lists. For half the *Ss*, one set of four pairs on the second list served as the synonym-pairs; for the other half of the *Ss*, these were the control-pairs.

Words were presented with a Dunning Animatic filmstrip-projector onto a milk-glass screen 10 in. square, so arranged that *S* could see neither *E* nor th eapparatus.

Following the anticipatory interval, the stimulus and response were always exposed together for 2 sec. Twenty anticipatory trials were given on the first list, and 16 on the second list. No mention was made that there was any relationship between any of the words in the two lists, nor of the change in anticipatory interval.

The experiment was replicated exactly, using college students as *Ss*. In this study, 45 *Ss*, 15 per group, were enrolled in the summer session at the University of California, Los Angeles, taking courses in psychology requiring participation in experiments.

Results: First list. The mean number of correct responses to the stimulus-syllables in the first list which were to be used in the second list is shown in Table I. The data on the other four pairs, which were used as fillers, will not be presented. Analysis of variance performed on these data resulted in an *F* of less than 1.00 for the high school *Ss*, but a significant *F* of 4.01 for the college *Ss*. This was primarily due to the 2-sec. group which had a few *Ss* who performed very poorly on the first list. Discussion of the difficulties brought about by this difference will be deferred until later.

TABLE I

MEAN NUMBERS OF CORRECT RESPONSES IN LIST 1 TO STIMULI USED IN LIST 2

Ss	1 sec.		2 sec.		4 sec.	
	M	σ_m	*M*	σ_m	*M*	σ_m
Highschool	47.40	7.02	47.25	3.58	41.75	2.69
College	61.17	2.41	45.63	5.72	56.75	2.98

Second list. The mean number of correct responses for both groups of *Ss* in blocks of two trials is shown in Table II. The results for the high school *Ss* showed no regular differences in performance between synonym-pairs and control-pairs except on the first trial. The exception to this was the 1-sec. group where there are some apparent positive transfer-effects through Trials 3.

---

[5] The authors wish to thank Mr. R. W. Dickinson and Mr. C. G. Warfel of Bonita High School, La Verne, Calif., for providing *Ss* for this study.

Testing these findings statistically presents some problems. Certain considerations made it advisable to consider only the data from the first anticipatory trial. Even though it may be expected that errors of Type I may be slightly inflated by this procedure, there is some *a priori* rationale for selection of these data since any transfer-effects would be expected to occur at that point if at all. It might seem that the design of the experiment conforms to the model for an analysis of variance of repeated measurements. It is difficult, however, to justify analysis of variance on these data since the scores can only be the integral values from zero to four, and under some conditions actually took on values of 0, 1, or 2. Since the effects of these kinds of restrictions on the analysis of variance are unknown, it was decided to use a nonparametric test, Friedman's Chi-square.[6] This test is probably a somewhat conservative one despite the fact that the results for each anticipatory interval must be tested separately.

TABLE II

Mean Numbers of Correct Responses by Blocks of Two Trials

	Rate		Blocks of trials								
			1	1–2	3–4	5–6	7–8	9–10	11–12	13–14	15–16
High school *S*s	1:2	Syn.	1.0	1.3	1.7	2.0	2.3	2.5	2.6	2.7	2.9
		Cont.	.4	.6	1.7	1.9	2.3	2.5	2.8	2.9	3.0
	2:2	Syn.	1.5	1.7	2.2	2.9	3.1	3.3	3.5	3.7	3.9
		Cont.	.4	1.0	2.2	2.6	3.2	3.3	3.6	3.6	3.8
	4:2	Syn.	1.5	1.5	2.5	3.3	3.5	3.8	3.8	3.9	3.8
		Cont.	.8	1.3	3.0	3.3	3.8	3.7	3.8	3.8	3.9
College *S*s	1:2	Syn.	.7	1.0	2.6	3.1	3.1	3.3	3.5	3.5	3.5
		Cont.	.5	.8	2.3	2.9	3.3	3.2	3.7	3.6	2.8
	2:2	Syn.	1.1	1.5	2.6	3.0	3.4	3.5	3.4	3.6	3.6
		Cont.	.8	1.3	2.6	3.0	3.6	3.5	3.5	3.6	3.7
	4:2	Syn.	2.8	3.1	3.7	3.8	3.8	3.9	3.9	4.0	4.0
		Cont.	1.8	2.4	3.6	3.7	3.8	4.0	4.0	4.0	4.0

TABLE III

Friedman Chi-Square Tests for All Conditions

Rate	High school *S*s		College *S*s	
	Chi-square	*p*	Chi-square	*p*
1:2	1.80	< .20	.03	< .80
2:2	6.05	< .02	1.03	< .05
4:2	2.45	< .20	6.90	< .01

The results of this analysis are presented in Table III. Significant positive transfer was obtained only with the 2-sec. group, although the direction of the difference favored the synonym-pairs in all instances. It might also be noted that increasing the anticipation-interval from 1 to 2 sec. increased the number of synonym-responses but had little effect on the number of control-responses, while increasing the interval from 2 to 4 sec. resulted only in an increase of control-responses.

---

[6] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, 166.

The results from the college *S*s, also shown in Tables II and III, are more clear-cut. Again consistent effects of synonymity are apparent only on Trial 1, although the small effects in the 1-sec. condition persisted for two more trials. Statistically significant positive transfer was obtained only at the 4-sec. interval which might be expected on the basis of the study of Barnes and Underwood.[7]

A clear interpretation of these results is marred by the initial differences between the groups. Although these initial differences in performance on List 1 suggest an adjustment by analysis of covariance, this analysis was not performed. The rank-order correlations between number of correct responses on the first list and amount of transfer were only —0.16, 0.12, and —0.09 for the three groups. Since it is not logically permissible to assume that degree of learning of the first list does not affect transfer, it can only be stated that despite differences in performance on the first list, amount of transfer does not seem to be related to these differences within the range of differences obtained here. Essentially, then, the groups differ on a characteristic which is not relevant to the amount of transfer. Furthermore, only 2 *S*s in the 2-sec. condition failed to reach at least one perfect trial during the 20 trials on the first list. Hence, it is felt that the lower degree of first-task learning is not responsible for the lack of evidence of positive transfer in this group.

Analysis of errors. Another way of looking at these data is to consider the events which occurred when *S* failed to make a correct response on the second list. Three possible events could occur: intrusions (responses from the first list); intralist errors (responses from the second list); or omissions (failures to respond). The frequency of occurrence of these responses is shown in Table IV, where it is apparent that the pattern of errors differed for the college *S*s and highschool *S*s. Although the frequency of both intrusions and intralist errors is highest in the 2-sec. condition for both highschool and college *S*s, the frequency of intralist errors is lowest for the 1-sec. condition for the highschool *S*s, while it is lowest in the 4-sec. condition for the college *S*s. Of the total errors made at the 4-sec. interval, intralist errors account for about one-third for both high school and college *S*s. Thus, the failure of the high school *S*s, to show significant positive transfer at the long interval cannot be attributed to intralist interference.

---

[7] Barnes and Underwood, *op. cit.*, 97-105.

Discussion. The principal findings of this study were: (1) An increase in positive transfer due to synonymity at longer anticipation-intervals. This was most apparent in the college *S*s. (2) No significant transfer-effects after Trial 1 on the second list. (3) More overt errors with a 2-sec. anticipation-interval as compared with longer and shorter intervals.

Only in the college *S*s at the 4-sec. anticipatory time were the positive transfer-effects due to synonymity really consistent. Of the 20 *S*s in each of the high school groups, 8 in the 1-sec. group, 12 in the 2-sec. group, and 11 in the 4-sec. group showed positive transfer. Of the 15 college *S*s, 5 in the 1-sec., 7 in the 2-sec., and 11 in the 4-sec. group showed positive transfer. The safest conclusion from these data, then, is that positive transfer may result when there is a high degree of synonymity but that the effects are transitory and negligible unless *S* is allowed time to respond.

TABLE IV

FREQUENCY OF TYPES OF ERRORS

	Rate		Intrusions	Intralist	Omissions	Total Errors
High school *S*s	1:2	Syn.	10	40	407	457
		Cont.	6	54	523	583
	2:2	Syn.	17	89	208	314
		Cont.	14	118	122	254
	4:2	Syn.	3	82	154	239
		Cont.	0	81	159	230
College *S*s	1:2	Syn.	9	32	243	284
		Cont.	3	27	248	278
	2:2	Syn.	15	65	212	292
		Cont.	3	59	209	271
	4:2	Syn.	5	14	25	41
		Cont.	0	22	49	71

The role of the mixed list in determining these results is not clear. Intuitively one might expect that this kind of design would interfere with the mediational processes presumed to account for positive transfer. It would certainly make it difficult for *S* consciously to classify the responses as synonyms; he would not only have to recognize the stimulus-syllable as one that had occurred in the first list, but recall the response which had been learned to this syllable if he were to use this response as a mediator. Nevertheless, mediated positive transfer has been obtained in studies using mixed lists.[8]

The question might also be raised in light of the failure of the high school *S*s to show consistent positive transfer-effects as to whether they knew the meaning of the words. A short vocabulary-test administered to these *S*s following the experiment indicated that they did know the words although this does not insure that they recognized the relationship in meaning during the experiment or attempted to use this synonymity in learning the second list.

---

[8] K. J. Norcross, and C. C. Spiker, Effects of mediated associations on transfer in paired-associate learning, *J. exp. Psychol.*, 55, 1958, 129-134.

The appearance of intrusions does suggest that there is some interference from first-list responses so that positive transfer could have been masked by competition between responses. The 2-sec. anticipatory interval seems to be optimal for the overt occurrence of intrusions, although it is likely that the competition between responses is also present and may even be greater in the 1-sec. condition. The shorter anticipatory time may prevent *S* from giving either the correct response or the intrusion. It would, in fact, be possible to attribute all positive transfer-effects at the longer interval to the inhibition of competing responses from the first list rather than to a longer time for the occurrence of mediating responses. The possibility of such a mechanism has been suggested by Underwood.[9]

Regardless of the mechanism involved, however, it is clear that the positive transfer resulting from synonymity of responses is not great.

SUMMARY

Three groups of *S*s were run in an experiment on transfer in which nonsense-syllables were used as stimuli and adjectives as responses in a mixed list. Each *S* learned two lists, with the second list consisting of four new pairs and four pairs in which the stimuli were identical to those in the first list and the responses were synonyms of those in the first list. The first list was learned with a 2-sec. anticipatory interval and the second list at an interval of either 1, 2, or 4 sec. The experiment was replicated with high school *S*s and college *S*s.

Results showed significant positive transfer due to synonymity on the first trial, for the high school *S*s at the 2-sec. interval and for the college *S*s at the 4-sec. interval. Overt errors were highest with a 2-sec. interval.

---

[9] Underwood, The effect of successive interpolations on retroactive and proactive inhibition, *Psychol. Monog.*, 59, 1945, (No. 273) 1-33.

RHYTHMIC ACTIVITY AND THE PERCEPTION OF TIME

By BRUCE DENNER, SEYMOUR WAPNER, JOSEPH H. MCFARLAND,
and HEINZ WERNER, Clark University

This study is part of an attempt to examine the processses underlying the cognition of time from the viewpoint on an organismic theory, viz. the sensory-tonic field-theory of perception.[1] Though this theory has evolved from, and has been largely restricted in its application to, spatial properties of objects, its major tenet—that perception is a reflection of a relationship between organismic state and proximal stimulus—has applicability to problems of temporal experience.

A first step in the attempt at generalization requires reformulating the basic postulate in terms which have pertinence for temporal experience. This step is readily taken if one considers: (1) that our usual handling of the perception of objects represents a *special* case of abstracting properties of objects as if they exist independently of temporal change; (2) that experience is given in spatio-temporal terms, *i.e.* as changes in spatial properties of objects or relations between them in the course of time; and (3) that temporal features, such as the duration of an event, may be abstracted from the totality of experience, and treated as if they exist independently of changes in spatial properties or relations among objects.

In the present study, we make the assumption that apparent duration is a reflection of a relationship between *personal tempo* (manifested in the stable rhythm an individual adopts for a specific task) and *temporal extension of an event (duration)*. The former represents the organismic state in its temporal aspect and the latter the proximal stimulus in the fundamental postulate of sensory-tonic theory.

Temporal experience is studied here by experimentally manipulating personal tempo and evaluating its effects on apparent duration of an event.[2]

---

* Received for publication, August 21, 1961. This investigaton was supported in part by a PHS Research Grant M-348, from the National Institute of Mental Health, Public Health Service.

[1] See Heinz Werner and Seymour Wapner, Sensory-tonic field-theory of perception: Basic concepts and experiments, *Riv. Psicol.*, 4, 1956, 315-337; Wapner and Werner, *Perceptual Development*, 1957, 1-65, and Jonas Langer, Seymour Wapner, and Heinz Werner. The effect of danger upon the experience of time, this JOURNAL, 74, 1961, 94-97.

[2] That personal tempo, as reflected in finger-tapping, can be experimentally manipulated is evident from a study by Thomas Casey, The effect of intermittent rhythmic stimuli on personal tempo (unpublished Master's thesis, Clark University, 1959). He found that preferred finger-tapping rate changes significantly in concordance with rate of external rhythmic stimulation.

Our expectation is that change in tempo—induced through synchronization of finger-tapping with rhythmic stimuli at higher and lower rates than the preferred finger-tapping rate—will bring about changes in perception of temporal extent.

Method. The 18 *S*s used in the experiment proper were selected on the basis of preferred finger-tapping rate from a group of 84 *S*s. *S* was required to tap on a key. connected with a Cenco timer, for 60 sec. at a rate which seemed comfortable for him. For the 84 *S*s tested, preferred rate of finger-tapping ranged from 0.8 to 4.8 taps/sec. The 18 *S*s falling in the middle of the range were selected. The tapping rate for this group ranged from 1.8 to 2.4 taps/sec. Immediately prior to the introduction of the conditions of the main experiment, the preferred tapping rate for these *S*s was tested again. On this second test, their preferred tapping rate once again fell into the same range.[3]

In the experiment proper, the main task for *S* was to estimate an interval of time by the method of reproduction while tapping on a telegraph-key at one of three rates: (1) preferred, (2) higher than preferred, and (3) lower than preferred. Thus data of two kinds were obtained: (1) *S*'s adopted tapping rate under the various rates imposed; and (2) the physical interval of time demarcated by *S* during reproduction of the standard interval.

Judgments of duration were measured as follows. A standard-interval of 70 sec. was presented to the *S*. The onset of the interval was indicated by a flashing light which continued to flash until the 70-sec. standard-interval terminated. After a pause of 10 sec., the light commenced, which indicated to *S* the onset of the variable interval. When a period of time passed which, to *S*, appeared equal to the standard, he signaled *E*, marking the termination of the variable interval. As indicated above, the experimental conditions, under which *S* estimated the 70 sec. standard, consisted of variation of the tapping rate. Three rates were employed: (1) the preferred rate; (2) 1.3 taps/sec. above the *S*'s own preferred rate; and (3) 1.3 taps/sec. below it. Each of the three tapping rates during presentation of the standard interval was paired with each of these three tapping rates under reproduction, making a total of nine conditions.

The tapping rates were imposed in the following manner. *S* sat in a dimly lit room and fixated a 3-in. circle of light emitted from an electronic metronome which was calibrated for different speeds. His task was to synchronize his tapping on a telegraph-key with the flashing light from the metronome while estimating the interval. He was instructed not to use counting in estimating the durations.

The telegraph-key was connected to a Cenco timer that the tapping rate could be determined both during the presentation of the interval and during the reproduction. Before the experiment began, *S* was given practice at tapping slowly and quickly to facilitate his adoption of the non-preferred tapping rates.

The conditions of tapping rates were as follows for presentation of the standard and reproduction of it (adjustment of variable): (1) slow tapping during presentation—slow tapping during reproduction; (2) slow-preferred; (3) slow-fast;

---

[3] Evaluation of reliability with an independent group of 42 *S*s yielded a test-retest coefficient of 0.93.

(4) preferred-slow; (5) preferred-preferred; (6) preferred-fast; (7) fast-slow; (8) fast-preferred; and (9) fast-fast. One man and one woman were tested in each sequence of experimental conditions in a 9 × 9 Latin square design, making a total of 18 *S*s.

Results: (1) Tapping rates. The first question was whether the individuals were successful in adopting the non-preferred tapping rates. Separate analyses of variance were computed of the tapping rates during (1) presentation of the standard-interval and (2) reproduction of the interval. There are three trials for each of the three tapping rates in both analyses of variance. The data were analyzed as a combined latin-square and factorial design.[4]

If presentation is considered separately from reproduction, there are nine tapping measures, three for each of the three tapping rates. The mean value for the imposed topping rates for the three main conditions were 0.8 for the slow-tapping condition, 2.1 for preferred, and 3.4 for fast tapping. The differences among tapping rates are significant beyond the 1% level ($F = 455.4$ with 2/72 *df*). The mean adopted rates—0.7 for slow tapping, 2.1 for preferred, and 3.2 for fast tapping—differ significantly from one another.

Similarly, there are nine tapping measures for the period of reproduction. The mean adopted rates during reproduction were 0.7 for slow tapping, 1.9 for preferred, and 3.0 for fast tapping. The variation in tapping rates during reproduction of the interval is significant beyond the 1% level ($F = 307.3$ with 2/72 *df*). Thus, the adopted rates differ in keeping with the conditions imposed, both during presentation and reproduction.

There is a further question as to whether within conditions the adopted rate differs significantly from the imposed rate. This was evaluated as follows: the three trials for a particular tempo-condition were averaged for each *S*. This value (adopted tapping rate) was compared with imposed tapping rate for that *S* and the set of difference-scores (imposed minus adopted) was tested against the hypothesis of no difference. This procedure was carried out for each of the three tempo-conditions. There was no significant difference between imposed and adopted tapping rates for any condition (in either presentation or reproduction), except for the fast condition in reproduction. Thus, except for the latter condition, there was no reason to assume that the *S*s were not able to adopt the imposed rates, and here the difference was only 0.2 taps/sec.

---

[4] See L. S. Kogan, Variance designs in psychological research, *Psychol. Bull.*, 50, 1953, 25-26.

(2) Estimations of time. The measure used in this analysis, apparent duration, was the elapsed physical time perceived by *S* to be equivalent to the standard interval of 70 sec. Analysis of variance of this measure shows that tapping rates have a significant effect beyond the 1% level during presentation ($F = 67.4$ with 2/72 *df*), reproduction ($F = 66.9$ with 2/72 *df*) and their interaction ($F = 6.8$ with 4/72 *df*).

The findings in Table I may be first examined for the effects of variation of tapping rate on the estimates of time during presentation—independent of their effect during reproduction. Fast tapping during presentation is associated with the shortest elapsed physical time, 70.9 sec.; slow tapping is associated with the longest elapsed physical time, 84.0 sec.; and elapsed physical time while tapping at preferred rate falls between these extremes. In contrast, the over-all effect of tapping rate imposed during adjustment of the variable (reproduction) shows the opposite relationship, *i.e.* fast tapping during reproduction is associated with the longest elapsed physical time, 84.9 sec.; slow tapping is associated with shortest elapsed physical time, 72.1 sec.; and tapping at preferred rate falls between these extremes, 75.6 sec.

TABLE I

ELAPSED PHYSICAL TIME PERCEIVED TO BE EQUIVALENT TO 70-SEC. INTERVAL
(Entries are mean sec.*)

		Imposed tapping rate on presentation				
		slow	preferred	fast	Mean	*SD*
Imposed tapping rate on reproduction	Slow	75.5	73.1	67.8	72.1	11.7
	Preferred	81.7	74.3	70.9	75.6	14.2
	Fast	94.7	85.8	74.0	84.9	15.9
	Mean	84.0	77.8	70.9		
	SD	14.1	12.3	12.9		

* A difference between the means of 3.8 is significant at the 5% level.

On superficial grounds, these two sets of results look antithetical, but the contradiction is only apparent. The following analysis accounts for the differences in the time-scores and shows the results to be completely consistent. It must be noted that tempo affects the apparent duration of the standard during presentation, and it also affects the variable during reproduction. These two effects on the magnitude of the reproduced interval are opposite to each other. That is, if the effect of a particular tapping rate is to make the standard appear shorter, then, independent of tempo-conditions during reproduction, the reproduced interval will be shorter; on the other hand, independent of tempo-conditions during presentation of the standard, if a particular tapping rate which makes the interval appear shorter is presented during reproduction, then the reproduced interval will be longer. Thus, tempo has a consistent effect on the perception of time whether it be introduced during presentation or during reproduction of the standard: with a tapping rate relatively faster (slower) than preferred there is relative apparent shortening (lengthening) of the 70-sec. interval.

Given this analysis of the general effects, we may now turn to a comparison of particular conditions of testing. The results for the nine conditions present a coherent picture when one considers two aspects of the experiment: (1) with relatively faster (slower) tapping rate than preferred, there is relative apparent shortening (lengthening) of a physical interval, whether it be during presentation of the standard or during the reproduction; and (2) a tapping rate imposed during presentation of the standard changes the apparent duration of the standard, whereas a tapping rate imposed during reproduction of the standard interval affects apparent duration during measurement.

Obviously, if the rate of tapping during measurement in and of itself would tend to shorten objective duration, it would require a longer time to match a given duration of the standard. Thus, the relationship between the effects of tapping rate during presentation of standard and tapping rate during reproduction defines the outcome of the particular test-condition. For example, if there is a test-condition with fast tapping on presentation and slow tapping on reproduction, there is a relation between two effects: the apparent duration of the standard is shortened, and the apparent duration during measurement is lengthened; the consequence is that the most extreme experience of shortening of duration occurs under this condition. By contrast, we should obtain the greatest subjective lenthening when slow tapping on presentation of the standard is paired with fast tapping during measurement. Where the same rate of tapping obtained under presentation of the standard and under measurement, the estimates should fall between these extremes.

The means presented in Table I show agreement with this analysis. The test-conditions combining faster tapping under presentation with slower tapping under reproduction range from 67.8 to 73.1 sec.; the test-conditions combining slower tapping on presentation with faster tapping on reproduction range from 81.7 to 94.7 sec.; and the test-conditions with matching tapping rates on presentation and on reproduction are intermediate, *i.e.* they range from 74.0 to 75.5 sec.

The significant effects of tempo-conditions on time-estimations, and the consistency of the patterns of various combinations of conditions, permit the generalization that a context of a tapping rate faster than preferred leads to an apparent shortening of an objectively given time-interval whereas tapping at a rate slower than preferred leads to an apparent lengthening of time.

SUMMARY AND CONCLUSION

This study dealt with the estimation of time. The basic finding is that with a change in tempo (preferred rate of activity) there is a change in apparent duration. The findings support the generalization that in the context of a faster (slower) non-preferred rate of activity, a given physical event appears shorter (longer).

These results are consonant with the conceptualization of temporal experience within the framework of sensory-tonic theory. The basic postulate of this theory, that perception is a reflection of the relationship between organismic state and proximal stimulus, is here generalized to temporal experience: apparent duration is a reflection of the relationship between personal tempo (temporal aspect of organismic state) and temporal extent (proximal stimulus).

TASKS, TASK-INTENSITY, AND DOLLARS DEMANDED FOR PERFORMANCE

By EDWARD A. BILODEAU and JAMES W. JOHNSON, Tulane University

Several years ago, Thorndike conducted an experiment in which price in dollars was the dependent variable.[1] *S* was asked to respond with the amount of money necessary for him to have the little finger of one hand cut off, become totally blind, become insane, and so forth. The results of this famous questionary are now quite out of date.

Many improvements upon Thorndike's method are now possible. Most of his questions pertain to mayhem and mutilation, and are not representative of the population of everyday human activities. There are other ways this study might be improved. The major limitation the writers find, however, is that Thorndike's questions were scaled by *S*'s responses. That is, there is no measure of the stimulus-variable taken independently of *S*. The items, as in so many other studies in the field of testing, were not used as variables from which functional relationships could be derived. In the present study, the S- and R-variables are independent and the inquiry is directed to the relationships between them.

The independent variable, task intensity, was varied quantitatively by using four *Levels* of each question, *e.g.* have two, four, six, or twelve children. Treating the independent variable in this way makes it possible to compare dollar responses both within and between questions. An *S* was given but one Level per question, so that the analysis within questions (between Levels) was based upon independent groups. The design also allowed for analyzing the relationship of the price *S* set to the serial position of the question. The items were selected from a population involving issues and tasks related to such diverse activities as politics, ethics, sex, family, work, and education.

It was hypothesized that the major effects of question and of Levels would lie in measures of individual differences or variability. Individual differences were expected to vary lawfully with stimulus-variables.

Method. A pool of over 200 questions was assembled from which 40 were selected for the final version of the questionary. The criteria used were clarity,

---

* Received for publication December 28, 1960.

[1] E. L. Thorndike, Valuations of certain pains, deprivations, and frustrations, *J. genet. Psychol.*, 51, 1937, 227-239.

interest, relevance to our time, and heterogeneity of action. Next, each question was subdivided into four numerical Levels producing 160 items. To serve in some branch of the Armed Forces for 1, 2, 4, or 8 years, illustrates the fractionation of one question into its four Levels.

Each *S* saw and answered 40 of the 160 items. Only one of the four Levels of a question was represented on *S*'s questionary. The order of the questions was established by a table of random numbers. Eight forms were used, four of which were written in the reversed order for the investigation of a sequence-effect. The four Levels were each represented 10 times per form. Ten Latin squares were used to determine the serial position of the Levels.

Procedure. The questionary was administered to general psychology students during regularly scheduled classes—an equal number of forms being distributed to a class. The forms were so distributed that adjacent *S*s had different ones.

The test was covered by a set of written instructions which included the following:

> Please think over the following instructions very carefully before turning to the questionary. Reread the instructions from time to time during the test if you need to.
>
> For what total amount of money in dollars, paid in cash and after taxes, would you do or undertake the following? You can spend the money in any way you desire. Fill in all the blanks as you go along, and do not go back to any question and change the answer. Do not race through the questions nor dabble along. If there is any item which you would pay to do, put the amount you would pay and circle it.

About 35 min. was required to answer the test. Twice, *E* interrupted and reminded the *S*s of the proper way to approach the questions. One of the more frequent questions asked of *E* was "What do I put if I would not perform the task under any circumstances?" The answer was: "Put *NA*, 'No amount.'"

Approximately 180 men were tested. This included four *S*s who did not wish to complete the test and one *S* who answered but two items. A few other *S*s were excluded at random in order to achieve an equal number of *S*s per item (42).

Results. Space does not permit our presenting the 40 questions and the results of each of their four levels.[2] Six of the questions are given in Table I as samples. The data are reported to the nearest dollar and with two significant digits retained. Any amounts over one million dollars are expressed by means of exponents. The six questions have been rank ordered in terms of the median for Level 1.

The columns in Table I represent Levels, the side headings the three quartiles. A comparison of the distances between quartiles will show a marked amount of skew; for this reason the results are analyzed by quartiles.

The results shown in Table 1 are typical of the entire battery. First, the median dollar values (Q_2) vary markedly with increasing Level. Second, individual differences are minor at the first Level, but increase rapidly from Level to Level. And third, the data become increasingly skewed with increasing Level.

---

[2] The complete questionary and the quartile results have been deposited with the American Documentation Institute.

TABLE I

NUMBER OF DOLLARS REQUIRED TO UNDERTAKE TASKS OF VARIOUS LEVELS OR INTENSITY

Question 1. Sleep _____ hours per night for the rest of your life and suffer no ill effects as a result of it?

Quartiles	2	4	8	12
Q_3	10,000	7,000	0	50,000
Q_2	*10**	0	0	180
Q_1	*10,000**	*100**	*25**	0

Question 4. Read _____ pages per day of your own choosing for one year?

Quartiles	10	30	90	270
Q_3	50	100	1,000	4,000
Q_2	0	1	100	500
Q_1	0	0	0	50

Question 15. Refrain from speaking for _____ day(s)?

Quartiles	1	5	25	625
Q_3	100	1,000	50,000	*NA*†
Q_2	50	100	2,500	1.2×10^6
Q_1	10	30	500	50,000

Question 25. Drive a _____ year old General Motors car for the rest of your life? The car is traded in every year on a different _____ year old car.

Quartiles	2	4	6	10
Q_3	10,000	20,000	50,000	100,000
Q_2	500	1,400	3,000	10,000
Q_1	0	0	100	100

Question 26. Be underweight by _____ pounds for the rest of your life?

Quartiles	10	20	30	60
Q_3	10,000	1,000,000	*NA*†	*NA*†
Q_2	880	11,000	750,000	5.0×10^{11}
Q_1	0	1,000	10,000	50,000

Question 33. Raise a child with an *IQ* of _____ ?

Quartiles	50	100	150	200
Q_3	1,000,000	250,000	1,000	100,000
Q_2	100,000	1,000	5	3,000
Q_1	10,000	0	0	0

* The italics in Question 1 mean that *S* is willing to pay.
† *NA* means that no amount would suffice.

It is obvious from the data presented in Table I that the full-scale relationship of dollars to Levels is not always a simple one, nor even monotonic. The over-all function, as in Question 33, resembles a U. Such functions are not adequately sampled by four Levels, and the Levels employed in the present study generally attempted to work on one side or the other of these non-monotonic curves.

Fig. 1 shows the medians for the two questions, "Be alone for —— days in a very small, sparsely furnished room" and the question "Be alone for —— day(s) in a very small, sparsely furnished room aboard a rocket ship going to Mars." In terms of dollars, the results make it obvious that *S* would much rather be confined to a small room (without indication of where) than to the same type of room aboard a rocket ship. The disparity between the medians (Q_2) is greatest at the lower Levels, and at Level 4, the amount required is exactly the same. At Levels beyond 1000 days (a log of 3) it would appear that the rocket ride would be preferred. The variability—not shown here in Fig. 1—is greater for the rocket ride, *especially* when the lower Levels are compared. The question used in Fig. 1 is beside the point. Fig. 1 was constructed to reëmphasize the virtues of scaling both independent and dependent variables in questionaries.

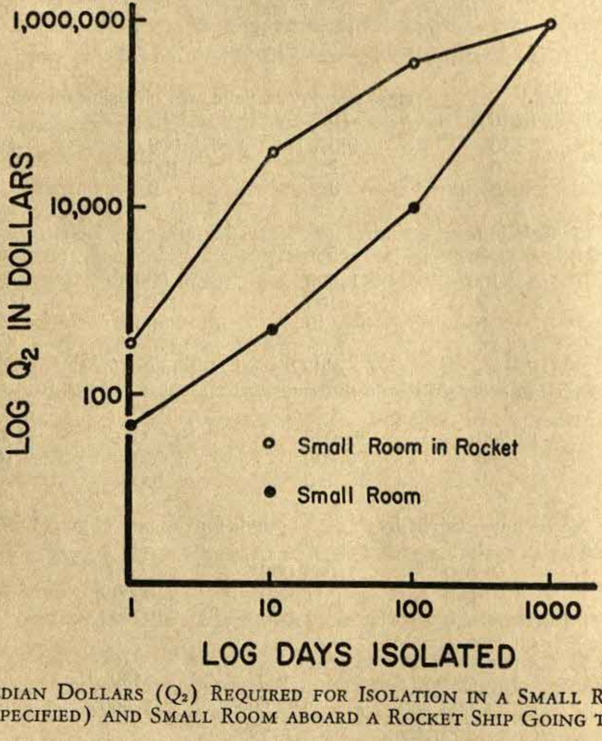


FIG. 1. MEDIAN DOLLARS (Q_2) REQUIRED FOR ISOLATION IN A SMALL ROOM (NO PLACE SPECIFIED) AND SMALL ROOM ABOARD A ROCKET SHIP GOING TO MARS

A check for a sequential effect was made. The analysis indicated that the amounts required tended to rise as *S* answered successive questions. This analysis was made by computing the difference between medians when a question appeared in the first and second halves of questionaries that were mirror images, and when expressed as a function of the degree of separation of the positions of the questions on the two forms.

Discussion. No lengthy discussion of the results is called for. This experiment was not designed to suggest why individuals and groups answer as they do. Rather, its purpose was to show that powerful relationships are at work when both independent and dependent variables are appropriately scaled. In this sense, the content or the subject matter of the battery was trivial. Of the several relationships indicated in the section on Results, the writers wish to call special attention to the increasing differences between individuals with increasing Level (an S–R type of relationship) and to the covariation of variability (Q_3–Q_1) and central tendency (Q_2), which is an R–R type of relationship.

There may, of course, be a legitimate interest in the present questions and answers *per se.* Two applications are obvious. First, there may be some value in assigning the questions to a stimulus-class defining incentive-motivation. The dependent variable, dollars demanded, has almost always been interpreted as a response of this class. Thus, it is possible to think of the antecedents and consequences of motivations as adequately handled by the present technique. The second possibility is to use the present items and answers as material for lectures in elementary psychology. Thorndike's work has often been used to show one thing or another about what people will do, and so may the present findings.

The obtained sequence-effect, though expected, was alarmingly powerful. In psychophysics, difference-limens, constant errors, and the like, depend upon *in* and *out* directions of stimulus-movement. As a consequence, averaging over both directions of movement is the recommended, standard psychophysical procedure. In paper and pencil testing, there is seldom control for direction of movement. The present findings do no more than suggest that a systematic inquiry be directed toward determining the conditions under which question-order becomes a relevant testing variable.

SUMMARY

This study developed further a task used by Thorndike some years ago. *S* was asked for how much money he would be willing to undertake

various tasks; this was the dependent variable. Each question had four intensities, or Levels, as the independent stimulus-variable. An *S* saw but one Level per question. There were eight forms of the battery, four being mirror images.

The data produced confirmed the usefulness of this method. In most cases, there was a direct relationship between the median amount of money demanded (Q_2) and the Level of the item. Individual differences (Q_3–Q_1) were very much related to Level and to the median (Q_2). The distributions of scores by Level generally showed positive skew which changed markedly with successive Levels. The data indicated a general tendency for the median dollars to rise as *S* answered successive questions.

THE ROLE OF AGE AND VERBALIZATION IN PROBABILITY-LEARNING

By HAROLD W. STEVENSON, University of Minnesota, and
MORTON W. WEIR, University of Illinois

Numerous studies have indicated that in tasks involving probability-learning the frequency with which *S*s choose the reinforcing stimulus is a function of the schedule of reinforcement employed. Comments of *S*s after being tested, however, indicate that often complex types of behavior occur which are not apparent in the usual measures of performance. For example, *S*s often indicate that they were responding on the basis of complex hypotheses or that they knew the stimulus which yielded the higher frequency of reinforcement even though they did not select it consistently. Typically, such verbal responses are made implicitly during performance and it is impossible for *E* to determine the types and frequencies of such responses or whether such responses have a significant influence on behavior.

In the present study, some *S*s were required to make overt verbal responses prior to each motor response. This was accomplished in two ways. First, the *S*s were tested in pairs with the instruction that they must reach a joint decision regarding each response. To insure that verbalization occurred, a second *E* was present with half of the pairs to monitor performance. Other *S*s were tested individually with the instruction that they were to verbalize to the second *E* the basis of each response. In an additional group, the *S*s were tested in the usual fashion alone and without a second *E*. The question was whether performance would differ when the *S*s were tested alone or in pairs and when they were or were not required to verbalize the basis of their response.

A second question was also explored. It is possible that differences among the conditions might not arise with the older *S*s since covert verbalizations may occur frequently and their performance may be as strongly affected by covert as overt verbalization. The younger *S*s may be less likely to make such frequent covert verbalizations and, because of less experience in utilizing such verbal stimuli as a basis of response, may respond quite differently when they are forced to verbalize from when they are

---

* Received for publication July 24, 1961. This study was supported by a grant from the National Science Foundation (Grant 3280) to the first author.

not. Consequently, *S*s at three chronological age-levels, 12, 15, and 18, were tested in each of the four conditions.

Method: (1) Subjects. The *S*s were 78 twelve-year-olds enrolled in the elementary school and 78 fifteen-year-olds enrolled in the junior high school of the University of Minnesota, and 78 eighteen-year-old students enrolled in the introductory classes in psychology at the University of Texas. The last group served in the study as a means of meeting the course requirement for participation in psychological experiments. The first two groups of *S*s were selected on the basis of chronological age (*CA*) from class-rolls. The *S*s were assigned to the conditions at random; and both the same and different sexes were tested in pairs.

(2) Apparatus. The apparatus has been described in detail elsewhere.[1] Essentially, it consisted of a yellow panel with a horizontal row of three black knobs. Above the knobs was a signal-light, and below the knobs was an aperture through which a marble fell into a plastic container. Surrounding the apparatus was a 4 × 6-ft. frame of black cloth to shield *E* from *S*'s view. A microphone was concealed behind the cloth and at the level of *S*'s head. All of the *S*s' verbalizations were recorded by a concealed tape recorder. Marbles were used as reinforcing stimuli.

Procedure. The *S*s at each level of *CA* were divided at random into four groups. Two groups of *S*s were tested individually and the other two groups were tested in pairs. It became apparent during early testing that pairs of *S*s often communicated with each other by means of inaudible whispers and gestures which could not be recorded on the tape. Consequently, both pairs of *S*s and individual *S*s were tested with a second *E* present. The second *E* sat adjacent to and behind the *S*s and instructed them to explain the basis of each choice to him. If they did not, *E* asked such questions as "On what basis are you choosing now?" Care was taken that *E*'s questions did not systematically follow reinforced or non-reinforced trials.

E brought *S* (or the *S*s) into the testing room and seated him (them) in front of the apparatus. *E* directed *S*'s attention to the knobs and the signal-light, then explained that each time the signal-light came on *S* was to press one of the knobs; if his choice was correct, a marble would be ejected into the box. *S* was instructed to get as many marbles as possible and that the signal-light would go off after each response, regardless of whether the response was correct.

When pairs of *S*s were tested, *E* added that the *S*s were to discuss every response with each other. In the appropriate groups, E_1 introduced E_2 and said that he was interested in the basis of their responses and that as long as they were discussing their choices loudly enough for him to hear them, he would not question them. When single *S*s were tested with E_2, E_1 instructed them to explain the basis of each choice to E_2 before pushing a knob.

A contingent procedure of reinforcement was employed, whereby the *S*s were reinforced for 33% of their choices of the correct stimulus. For each *S* (or pair of *S*s) one of the three knobs was designated as the 'correct' knob. The correct knob was determined by a prearranged sequence, and an equal number of *S*s was tested with each knob. The other two knobs never yielded reinforcement. The schedule of reinforcement was randomized in blocks of nine trials. Thirty such blocks were constructed, and each *S* was started at random on any of the 30 blocks to eliminate sequential effects in the schedule of reinforcement. The sequence of reinforcements was randomly determined with two restrictions: no more than two successive choices of the correct knob yielded reinforcement and no more than three consecutive choices of the correct knob were not reinforced. Each *S* (or pair of *S*s) was tested for 90 trials. Latencies of responses were measured for the oldest group of *S*s by means of an electric timer wired so that the onset of the signal-light activated the timer and *S*'s pushing one of the knobs stopped it.

---

[1] H. W. Stevenson and E. F. Zigler, Probability learning in children, *J. exp. Psychol.*, 56, 1958, 175-191.

Results. The measure used in the analysis of results was the incidence of choice of the correct knob. The mean number of responses made to the reinforcing knob by each group and at each age-level is presented in Fig. 1 for each block of nine trials. An analysis of variance of these data reveals a significant difference in performance at the three age-levels ($F = 25.05$, $df = 2/144$, $p < 0.001$). The incidence of correct response increases with increasing chronological age. Performance in the four conditions does not differ significantly ($F = 1.82$, $df = 3/144$, $p > 0.05$), but the interaction between Condition and Age is significant ($F = 2.80$, $df = 6/144$, $p < 0.05$). The basis of this interaction appears to lie in the greater difference in performance among conditions at the two older as compared to the younger age-levels. As is apparent in Fig. 1, there is a significant tendency for the incidence of correct response to increase across the blocks of trials ($F = 52.28$, $df = 9/1296$, $p < 0.001$). The triple interaction of Age by Condition by Trials is significant ($F = 1.88$, $df = 54/1296$, $p < 0.01$).

The *S*s at all age-levels were able to verbalize the basis of their responses, and E_2 had to question the *S*s relatively infrequently. To analyze the *S*s' verbal responses, some form of categorization was required. The transcribed records were carefully read to determine the different types of verbal responses made by the *S*s. A total of 22 empirically defined categories evolved from this procedure. These were eventually reduced to seven broader categories. Most of these categories did not involve interpretation, but were distinguished by the *S*s' statements differentiating the category. The categories and an example of each are given below.

(1) Schedule contingent on E. The *S*s stated that *E* was changing the reinforcement-schedule, was changing the reinforcing stimulus, or was not loading the apparatus. Example: "I'm sure that the marble comes out of that one, but I'm still afraid that he might change it."

(2) Contingent upon structure. The *S*s stated that they were responding on the assumption that correct response was dependent upon learning some pattern or sequence of responses. Example: "Well, the point is to get as many marbles as we can. So there has to be a pattern back there somewhere. So let's just press and

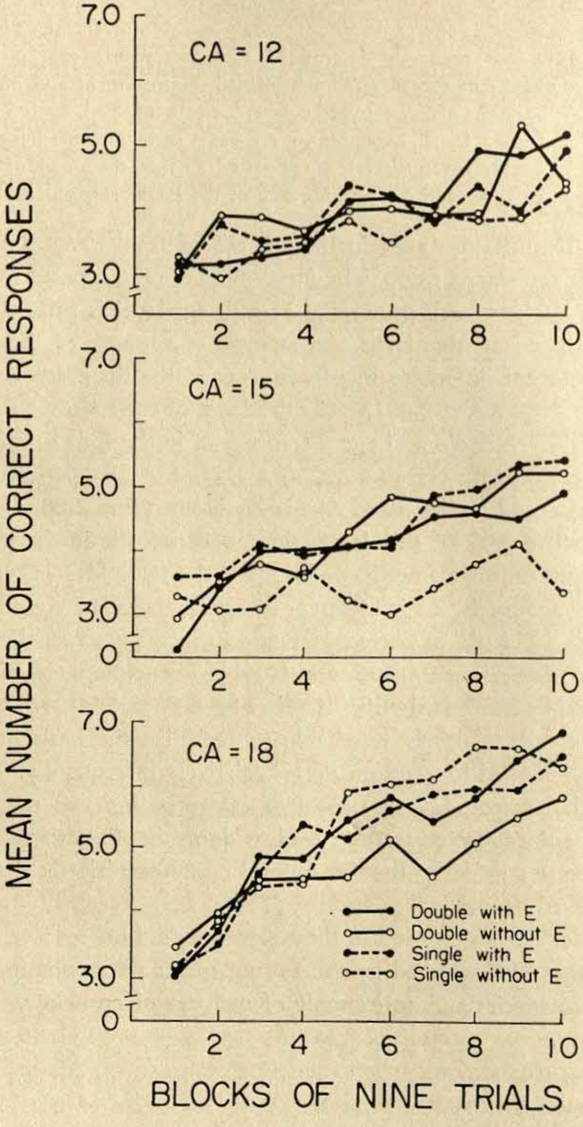


FIG. 1. MEAN NUMBER OF CORRECT RESPONSES IN BLOCKS OF NINE TRIALS AS A FUNCTION OF AGE AND EXPERIMENTAL CONDITION

try to find the pattern. Just press the different ones until we find the sequence they go in it. They've got to have some kind of order."

(3) Contingent upon previous responses. The *Ss* stated that a stimulus should be chosen because it had not been chosen recently or because the present choice had been selected too frequently and should be abandoned. After a reinforced trial, *Ss* said that a new response would be required since the prior response had been reinforced. In general, any statement that the present choice was determined by prior choices was included in this category. Example: "I'll push this one because I haven't pushed it for a while."

(4) Affective responses. The *Ss* stated that they were selecting a stimulus because they liked it, wanted to, or felt that it would be correct. Example: "I just have a feeling that it will be there. There isn't any other reason."

(5) Response to irrelevant stimuli. The *Ss* based their choices on stimuli which had no logical relationship to the task. Example: "Let's push No. 2 because there are two of us."

(6) Guessing. The *Ss* stated that they were not approaching the task systematically. Example: "We're just jumping around until we find out how it works."

(7) Do not know. The *Ss* stated that they did not know why they made a particular choice. Example: "I don't have any reason for pushing this one."

The measure used in the statistical analysis of the verbalizations was whether or not *S* made at least one statement under a particular category. Differences as a function of age were explored first. There were no differences associated with age for the first category. Twelve *Ss* at *CA* 12 and 14 *Ss* at *CA* 18 stated that their response was contingent in some way upon *E*. All *Ss* at each level of *CA* made at least one statement that their response was contingent upon structure (Category 2), and that their response was contingent upon previous responses (Category 3). The number of *Ss* stating an affective basis for response decreased from 10 at *CA* 12 to 5 at *CA* 18, but the difference is not significant ($\chi^2 = 3.32$, $df = 2$, $p < 0.10$). There was a decrease in the number of *Ss* stating irrelevant bases for their responses. At *CA* 12, 23 *Ss* made at least one statement; at *CA* 15, 14 *Ss*; and at *CA* 18, 10 *Ss*. The difference is highly significant ($\chi^2 = 14.24$, $df = 2$, $p < 0.001$). There was a significant decrease in the number of *Ss* who guessed from 17 *Ss* at *CA* 12 to 8 *Ss* at *CA* 18 ($\chi^2 = 6.50$, $df = 2$, $p < 0.02$). There was also a significant decrease in the number of *Ss* stating that they did not know the basis of their responses from 20 at *CA* 12 to 11 at *CA* 18 ($\chi^2 = 6.82$, $df = 2$, $p < 0.01$).

When the data are analyzed according to whether *Ss* were tested individually or in pairs, only the difference in the number of *Ss* stating that response was contingent upon *E* was significant. There were 12 *Ss* who made this statement when they were tested alone, while 27 *Ss* made this statement when tested in pairs ($\chi^2 = 11.54$, $df = 1$, $p < 0.001$).

The categories containing the greater proportions of all remarks were Category 2 (37% of all remarks) and Category 3 (36% of all remarks). The most common explanations of response were, therefore, that response was dependent upon structure or upon previous responses.

Information concerning the amount of time devoted to discussion and decision is reflected in the latencies of response, which were collected for the oldest group of *S*s. The mean latencies after *S*s received a reinforcement were greater than after non-reinforcement ($t = 2.94$, $p < 0.01$). An analysis of variance of the *S*s' mean latencies following a reinforced trial revealed a significantly longer latency when E_2 was present ($F = 9.66$, $df = 1/48$, $p < 0.01$) and when pairs of *S*s were tested ($F = 23.41$, $df = 1/48$, $p < 0.001$). Similar results are obtained in an analysis of variance of the latencies following a non-reinforced trial. The presence of E_2 increased decision-time ($F = 9.18$, $df = 1/48$, $p < 0.01$), and pairs of *S*s took longer to decide than did single *S*s ($F = 32.86$, $df = 1/48$, $p < 0.001$). The interaction between the two variables was also significant ($F = 8.27$, $df = 1/48$, $p < 0.01$).

Discussion. The four experimental conditions failed to produce significant differences in performance. It must be concluded, therefore, that under these particular testing conditions, two *S*s do not perform differently from a single *S*, and *S*s who are forced to verbalize the basis of their responses to a second *E* do not differ in their choice-behavior from *S*s who do not verbalize explicitly.

Only one *S* clearly stated the appropriate strategy for maximizing reinforcement, and only 8 *S*s (or pairs of *S*s) maximized their reinforcement by consistently choosing the reinforcing stimulus during the terminal 20 trials. The verbalizations were generally unrelated to the actual procedure of reinforcement. All *S*s had strong tendencies to assume that reinforcement was based on some pattern or sequence of responses and that it was determined by previous responses. There was no evidence to indicate that if *S*s did form more or different hypotheses as a result of having to verbalize the basis for their responses, such verbalizations had a significant effect on performance.

The younger *S*s appeared to have difficulty in explaining why they were making their particular choices, and often met the request to verbalize by saying that they did not know, were guessing, or by creating some irrelevant basis for response. Their irrelevant comments were often of a phenomenalistic nature, in that they would use any stimulus present, such as number of people, number of buttons, angle of light, shadows, etc., as the basis for their choice. The differences in performance as a function of age cannot be easily understood from the differences in verbalizations, nor can they be accounted for by current probabilistic models of choice-behavior. It is of interest that the performance of the 18-yr.-olds is most similar to that of 3-yr.-olds who were tested in a recent study comparing probability-learning in 3-, 5-, 7-, and 9-yr.-olds.[2] The 3-yr.-olds and the 18-yr.-olds have the highest frequency of choice of the reinforcing stimulus of all the groups tested in the present and in the earlier study. A curvilinear relationship between chronological age and frequency of choice of the correct stimulus has also been reported in another study of discriminative learning comparing the performance of *S*s from 4 through 18 yr. of age.[3]

SUMMARY

The experiment investigated the performance of *S*s at *CA*s 12, 15, and 18 yr. in a task in which one response yielded 33% reinforcement and the other two responses never yielded reinforcement. Four conditions were employed: *S*s were tested alone or in pairs, and with or without a second *E* monitoring their performance. In the latter two conditions the *S*s were told to verbalize the basis of their responses or to reach joint decisions concerning their responses. Performance under the four conditions did not differ significantly, but a significant tendency for older *S*s to make more frequent choices of the reinforcing stimulus was found. There was a significant interaction between Condition and Age. Differences in verbalizations were explored, and significant decreases in the numbers of *S*s who stated that they were guessing, did not know, or used irrelevant bases for response were found with increasing age. More of the *S*s tested in pairs were found to state some behavior of *E* as the basis of response than *S*s tested alone. It is concluded that testing *S*s alone or in pairs, or with or without a second *E*, does not significantly influence performance in the type of task employed.

---

[2] H. W. Stevenson and M. W. Weir, Variables affecting children's performance in a probability learning task, *J. exp. Psychol.*, 57, 1959, 403-412.

[3] H. W. Stevenson, Ira Iscoe, and Claudia McConnell, A developmental study of transposition, *ibid.*, 49, 1955, 278-280.

EVIDENCE FOR A RELATIONSHIP BETWEEN FIGURAL AFTER-EFFECTS AND AFTER-IMAGES

By ROBERT F. TERWILLIGER, Rutgers, The State University

Köhler and Wallach report noticing after-images of inspection-figures (*I*-figures) during the prolonged fixation of the *I*-figure which is necessary to produce a figural after-effect (*FAE*).[1] It is conceivable that the distortion in the test-figure (*T* figure) is dependent on the presence of this after-image. Both Köhler and Wallach and Gibson have, however, demonstrated that in a sense the *FAE* is independent of the eye stimulated.[2] That is, *S* can fixate the *I*-figure with one eye and the *T*-figure with the other eye, and a *FAE* will still be reported. Köhler and others have used this sort of evidence to argue that any theory of retinal fatigue could not account for the distortion of the *T*-figure in the *FAE*-situation.[3] Retinal fatigue, hence the after-image, is restricted to a particular place on a particular retina. It is, therefore, impossible by this theory to explain the distortion of a *T*-figure presented to the right eye when any after-image of the *I*-figure is in the left eye. It has been demonstrated, however, that after-images show interocular transfer.[4] Evidence therefore, exists proving Köhler and Wallach wrong on this issue of transfer of retinal fatigue.

Before, one could use the theory of retinal fatigue to explain *FAE*, however, other features besides the interocular transfer of after-images must be shown. There are several sets of results in the literature which demonstrate that it is the retinal size of the stimuli which determines the direction of distortion of *FAE*, that this law of retinal size resembles Emmert's law, and that there is no discernible *FAE* when retinal sizes of

---

* Received for publication June 12, 1961. This study was made in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Stanford University. The author wishes to thank Professors D. H. Lawrence, Quinn McNemar, and A. H. Hastorf of Stanford and Dr. J. S. Adams of The General Electric Company for their assistance.

[1] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

[2] Köhler and Wallach, *op. cit.*, 271; J. J. Gibson, Adaptation, after-effect, and contrast in the perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1-31.

[3] Köhler and Wallach, *op. cit.*, 340; Wolfgang Köhler and D. A. Emery, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159-201.

[4] A. K. Hansen, "After-image transfer test" in anomalous retinal correspondence, *Arch. Ophthal.*, 32, 1954, 369-374.

the *I*- and *T*-figures are equated.[5] All the above argue strongly for a retinal interpretation of *FAE,* since both after-images and *FAE* appear to follow the same law of retinal size.

If it can be shown that *FAE* show interocular transfer and that this transfer follows a law of retinal size, and also that after-images show interocular transfer and Emmert's law is followed despite this transfer, then it is felt that the evidence would be quite strongly in favor of the following interpretation: *FAE* and after-images, since they follow the same retinal laws, are probably related to the same basic causal mechanism. The suggestion of Köhler and Wallach that retinal fatigue is not related to *FAE* is erroneous.

EXPERIMENT I: INTEROCULAR TRANSFER OF THE FIGURAL AFTER-EFFECT

Method and procedure. Through a monocular eyepiece the *Ss* viewed a table 10 ft. in length on which were stimulus-card holders at distances of 4 and 8 ft. from *S*. These holders were built to receive cards 30 in. in the horizontal direction by 10 in. in the vertical, and such that the cards could be slid easily in and out of them. The whole apparatus was illuminated from above by one 40-w. fluorescent bulb situated 50 in. above the table so that each card-holder received approximately the same amount of light, with the light evenly distributed over the surface of each card. The *I*-figure was a 2-in. circle either to the right or to the left of the fixation-point, which in all cases was in the center of the card, and placed 4 ft. from *S*. The *T*-figure consisted of two 2-in. circles, one on either side of the fixation-point, and placed 8 ft. from *S*. These figures were drawn in India ink on white illustration-board.[6]

Forty *Ss* were used, both men and women, all students in introductory psychology. Two measurements, 30 min. apart, were made on each *S*. In both, the eye used in fixating the *I*-figure and the side of the fixation-point on which the *I*-circle was presented were held constant for any given *S*. The measurements differed only in the eye to which the *T*-figure was presented. The *Ss* were randomly assigned to the various conditions so that half of them fixated the *I*-figure with the left eye and half with the right eye. Half of these had the *I*-circle on the right of the fixation-point and half on the left. Finally, half saw the *T*-figure with the right eye first and half with the left eye first. *S* was required to report one of the *T*-circles as larger. Responses of "equal" were not allowed.

Results. The results, presented in Table I, show the *FAE.* These are all statistically significant (exact $p = 0.005$ and 0.01 for left-eye *I*-conditions and 0.035 and < 0.001 for right-eye *I*-conditions) and in the direction predicted by a retinal hypothesis. The extent of the reported distortions was in all cases comparable to that reported in similar binocular *I*- and *T*-conditions (18 out of 20 in the retinally predicted direction).[7] It is necessary, to the hypothesis being tested, that transfer-conditions not be different from non-transfer conditions. McNemar's test for the significance of changes indicates that measurements of distortion made, when the inspection-eye and the test-eye were the same, were not different from measurements made when the inspection-eye and the test-eye were different ($X^2 = 0.308$, $p > 0.5$).[8]

---

[5] Julian Hochberg and M. E. Bitterman, Figural after-effects as a function of the retinal size of the inspection-figure, this JOURNAL, 64, 1951, 99-102; W. C. H. Prentice, The relation of distance to the apparent size of figural after-effects, *ibid.*, 63, 1950, 589-593; R. F. Terwilliger, Retinal size as a determiner of the direction of size distortion in figural after-effects, *Quart. J. exp. Psychol.*, 13, 1961, 209-217.

[6] Since the *T*-figure was retinally smaller than the *I*-figure, the retinal hypothesis would predict that *S* would report the left *T*-circle as smaller if the *I*-circle had been on the left of the fixation-point and the right *T*-figure as smaller if the *I*-circle had been on the right of the fixation-point.

TABLE I

FREQUENCY OF RESPONSES IN THE RETINALLY PREDICTED DIRECTION
(Based on 20 *S*s per column, measured once in each row)

Test	Inspection	
	Left-eye	Right-eye
Same eye	16	15
Other eye	16	18

The results support the claims of other investigators that there is interocular transfer of *FAE*.[9] Also, it has been shown that *FAE* follow a law of retinal size and that this is also subject to interocular transfer.

EXPERIMENT II: INTEROCULAR TRANSFER OF AFTER-IMAGES

Method and procedure. For the measurements of Emmert's law, a modification of the previously described apparatus was used. A card-holder was added 6 ft. from *S*. The eyepiece was removed and a head-rest substituted for it. To measure the size of after-images, each of the three card-holders was equipped with a set of calipers, which could be controlled by *S* from his position of observation. A removable light source, 3 ft. from *S*, consisted of a 1½-in. circular opening, on a level with and directly ahead of *S*'s eyes, covered with milk-glass and behind which was a Westinghouse No. 2 photoflood lamp. This apparatus was mounted on a large board which could be lifted out of the way to allow a clear view of the card-holders and the calipers. Once the after-image had been created, this apparatus was removed and *S* saw in one of the holders a white illustration-board card with a fixation-cross in the center. Both the creation of the after-image and the measurement of it were done in a room illuminated as in the previous study.

The order of measurement was the same for all *S*s. The first condition was fixating the light-source and measuring the after-image with the left eye, first with the card at 4 ft., then at 6 ft., and finally at 8 ft. The distances were repeated in this order for the left-eye fixating and right-eye measuring, right fixation and measurement, right fixation and left measurement, and binocular fixation and measurement, in that order. Before each measurement, the after-image was produced anew: When the 6 left-eye measurements were completed, *S* was given a 20-min. rest before starting the right-eye conditions and completing the experiment.

---

[7] Terwilliger, *op. cit.*, 214.

[8] Quinn McNemar, *Psychological Statistics*, (2nd ed.), 1955, 228-230.

[9] Köhler and Wallach, *op. cit.*, 271.

Four *S*s were used, three men and one woman. All had at least a B.A. in psychology or were doing graduate work in psychology. Only one had had previous experience making measurements of Emmert's law, and none was familiar with the hypothesis being tested. When an *S* was brought into the room, he was allowed to familiarize himself with the operation of the calipers. After fixation on the light-source for 7 sec., he was asked to adjust the calipers, which, at the beginning of each measurement were always at their most extreme separation, until the inside of the caliper-rods was tangent to the after-image. In the cases in which only one eye was used, *S* wore a pair of optician's eyeglass-frames into which an opaque black patch could easily be inserted and removed by *E* while they were being worn. In tests of eye-transfer, after the light-source had been fixated with one eye, *S* was asked to close both eyes. During this time, *E* switched the patch to the stimulated eye, exposing the unstimulated one. The *S* was then instructed to open his eyes and adjust the calipers as before. After each fixation and measurement, there was a 5-min. rest. If, at the end of this time, *S* still reported an after-image, additional rest was given until at least 1 min. had elapsed during which no after-image was seen.

Results. The data on after-images were treated by an analysis of variance which indicated that the particular eye-combination used had no effect on the perceived size of the after-image ($F = 1.33$) and that there were no significant interactions. Only the distance at which the measurement was made significantly influenced the apparent size of the after-image ($F = 1342.01, p < 0.01$).

Emmert's law enables one to predict the exact size of the measurement of the after-image which should be made at any given distance. An analysis of variance, described by Lewis,[10] for testing the goodness of fit of data to a predicted curve was used on the obtained results. The results used were the average of all *S*s over all conditions, which should be the most reliable estimate of the true mean of the measurements. The values obtained did not differ significantly from the ones predicted by Emmert's law ($F = 0.20$.)

The results of this study indicate that after-images show interocular transfer. The only condition affecting perceived size of the after-image was the distance of the surface from the *S*, as one would expect if Emmert's law was being followed. In addition, the results indicated that there was no difference between the obtained values and those which would be predicted by Emmert's law.

---

[10] Don Lewis, *Quantitative Methods in Psychology,* 1960, 372-379.

SUMMARY AND CONCLUSIONS

Evidence has been presented which is relevant to the problem of figural after-effects (*FAE*), and in particular to those size-distortions studied by Köhler and Wallach. Previous results have indicated that these size-distortions follow a law of visual angle in the same way that after-images follow Emmert's law, also a law of visual angle. The present studies show that both phenomena are subject to interocular transfer and that they both continue to follow a law of visual angle when transferred. The results suggest that these two phenomena may be quite closely related. Köhler and Wallach's rejection of a theory of *FAE* based on retinal fatigue or after-images on the ground that after-images do not show interocular transfer is totally invalid. Such transfer does take place, and thus *FAE* could be explained retinally. Such a state of affairs has not been proven, but it is worth considering.

EFFECTS OF MUSIC UPON *GSR* AND HEART-RATE

By George H. Zimny and Edward W. Weidenfeller,
Marquette University

The purpose of the present study is to investigate the effects of music judged to differ with respect to capacity to excite and to calm upon galvanic skin response (*GSR*) and heart-rate. The hypothesis being tested is that the responses to three pieces of music designated as exciting, neutral, and calming exist in that order on a scale of *GSR*, with the magnitude of the score inversely related to resistance, and on a scale of heart-rate, with the magnitude of the score directly related to the number of beats per minute. The results of previous studies on the physiological effects of music generally lend support to this hypothesis, at least with respect to *GSR*.[1] *GSR* and heart-rate were selected for study because they are assumed to be two physiological manifestations of emotional response.

METHOD

Subjects. The *S*s were 18 college students, 10 men and 8 women, randomly selected from a total of 78 summer school students enrolled in three lower-division courses.

Apparatus and materials. Selected 6-min. portions of three musical pieces—the final movement of Dvořák's *New World Symphony* (ML 5115, Columbia), Chopin's *Les Sylphides* (CL 741, Columbia), and Bach's *Air for the G String* (ML 5065, Columbia)—were judged by the authors to be exciting, neutral, and calming, respectively. Ratings of each piece obtained from 59 undergraduate students substantiated the authors' judgment.

A tape of the selected portion of each piece was played on a Wollensak tape-recorder over a high-fidelity loudspeaker. *GSR* was measured by means of a psychogalvanometer (Stoelting, Model No. 24207). The measure consists of reaction-units read from a dial located on the front of the galvanometer. In the present experiment, a change of 12.5 reaction-units from the zero dial-reading represented a change in resistance of 1000 ohms. This galvanometer does not permit the determi-

---

* Received for publication March 9, 1961. This research was supported in part by a grant from the Committee on Research of Marquette University.

[1] M. L. Phares, Analysis of musical appreciation by means of the psychogalvanic reflex technique, *J. exp. Psychol.*, 17, 1934, 119-140; R. E. Dreher, The relationship between verbal reports and the galvanic skin response, Unpublished Doctoral dissertation, Indiana University, 1947; D. S. Ellis and Gilbert Brighouse, The effects of music on respiration- and heart-rate, this JOURNAL, 65, 1952, 39-47; R. E. Henkin, The prediction of behavior response patterns to music, *J. Psychol.*, 44, 1957, 111-127; Werner Traxel and Gunhild Wrede, Hautwiderstandsänderungen bei Musikdarbietung, *Z. exp. angew. Psychol.*, 6, 1959, 293-309.

nation of a basal level. Before each of the three pieces of music was played to each *S*, the galvanometer was adjusted to match the unknown level of resistance of the *S*, resulting in a dial-reading of zero.

Heart-rate was measured by a photoelectric cardiotachometer developed by Zens and Mounts.[2] The cardiotachometer was attached to the ear-lobe and rate was determined by means of an electric counter.

Procedure. Each *S* was individually exposed to all three musical pieces. The six sequences of pieces were determined by complete counterbalancing, with each *S* assigned randomly to a sequence. *S* was seated alone in an easy chair in a soundproof booth. *E* and all recording and playing apparatus except the loudspeaker were located in an adjacent room from which *S* could be observed. The sequence of events for each *S* was as follows: instructions to *S* and connection of the galvanometer and cardiotachometer to the *S* (3 min.); pre-music measurement (1 min.); measurement with music playing (6 min.); post-music measurement (1 min.); rest-period outside of laboratory-booth (5 min.). The entire sequence was repeated two more times with a different musical piece. The *S*s were instructed to sit quietly and merely listen to the music. The number of reaction-units shown on the galvanometer and the number of heart-beats were recorded every 15 sec., yielding for each musical piece a total of 24 readings per *S* for each measure.

RESULTS

Heart-rate. Analysis of the results using heart-rate for each of the three musical pieces yielded no significant differences.

GSR. The mean *GSR* reaction-units for each successive minute for each piece of music were computed for each *S* from the four readings obtained during each minute. The slope of the line of best fit for these means was calculated for each *S*. The slopes were analyzed by a treatments × *S*s analysis of variance.[3]

The means and standard deviations of the 18 slopes for each piece of music were: Exciting, $M = 4.08$ and $SD = 2.02$; Neutral, $M = 0.29$ and $SD = 1.81$; Calming, $M = -0.64$ and $SD = 1.38$. The analysis of variance yielded an F of 32.34 ($P < 0.001$). Analysis of each pair of means resulted in a t of 6.04 ($P < 0.005$) between Exciting and Neutral, a t of 7.60 ($P < 0.05$) between Exciting and Calming, and a t of 1.51 ($0.10 > P > 0.05$) between Neutral and Calming.[4] All levels of significance are based on one-tailed tests. Only the last difference is not significant. Analysis of each pair of variances using the t-test of variances based on related scores yielded no significant differences.[5]

---

[2] Carl Zens and Frank Mounts, Determination of heart-rate during work, *Arch. indust. Health,* 7, 1958, 280-286.

[3] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education,* 1953, 156-158.

[4] Lindquist, *op. cit.,* 164-166.

[5] H. M. Walker and Joseph Lev, *Statistical Inference,* 1953, 190-191.

DISCUSSION

The hypothesis tested in the experiment was partially confirmed with respect to the *GSR* but not heart-rate. The response to the exciting music was highest on the *GSR*-scale, as predicted, but the response to the neutral and calming music did not differ. The *GSR*-response to the exciting music was a pronounced decrease in resistance, indicative of an increase in arousal, while the response to both the neutral and calming music remained unchanged throughout the playing of the music. On the assumption that the *GSR* constitutes a physiological manifestation of emotional arousal, the present study demonstrates that music judged by untrained persons to be exciting does have the expected emotional impact and that music judged to be neutral or calming has no emotional effect.

The lack of agreement between the findings based on the *GSR* and those based on heart-rate fits into the pattern of a growing body of research and theory. Both the *GSR* and heart-rate were used in the present study as indicators of emotional arousal. While Malmo states that the *GSR* and heart-rate can serve as indicators of arousal, some question has been raised as to their equivalence in this respect.[6] Lacey reports "directional fractionalization" of the response-pattern of heart-rate and palmar conductance under four different stimulus-conditions.[7] Davis and Buchwald report a contrast between heightened activity of skin and muscle and lowered action of circulatory and respiratory processes in response to pictorial stimuli.[8] In addition, the *GSR* and heart-rate can be differentiated on the basis of the schema of body-exterior vs. body-interior proposed by Fisher and Cleveland.[9] The present study provides additional support for this concept of differential functioning of the *GSR* and heart-rate and also indicates, on the practical side, the advisability of using the *GSR* rather than heart-rate in studies of the emotional effects of music.

SUMMARY

The hypothesis was tested that the responses to three pieces of music designated as exciting, neutral, and calming exist in that order on a scale of the *GSR* and of heart-rate. Selected 6-min. portions of pieces by

---

[6] R. B. Malmo, Anxiety and behavior arousal, *Psychol. Rev.*, 64, 1957, 276-287.

[7] J. I. Lacey, Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome, in E. A. Rubinstein and M. B. Parloff (eds.), *Research in Psychotherapy*, 1953, 160-208.

[8] R. C. Davis and A. M. Buchwald, An exploration of somatic response patterns: stimulus and sex differences, *J. comp. physiol. Psychol.*, 50, 1957, 44-52.

[9] Seymour Fisher and S. E. Cleveland, An approach to physiological reactivity in terms of a body-image schema, *Psychol. Rev.*, 64, 1957, 26-37.

Dvořák, Chopin, and Bach that were judged by college students to be exciting, neutral, and calming, respectively, were played in counterbalanced order to each of 18 randomly selected college students. Measures of the *GSR* and heart-rate were taken at intervals of 15 sec. The hypothesis was partially confirmed with respect to the *GSR* but not heart-rate. The *GSR*-response to the exciting music was a pronounced decrease in resistance (indicating increased emotional arousal), while the response to both the neutral and calming music remained unchanged throughout the playing of the music.

THE INFLUENCE OF COLOR ON DISCRIMINATION OF SWEETNESS AND SOURNESS IN PEAR-NECTAR

By ROSE MARIE PANGBORN and BRENDA HANSEN,
University of California, Davis

Although it is commonly accepted that people have color-flavor associations resulting from experience with natural and processed foods,[1] few controlled experiments have been reported on the direction and magnitude of this association.

Previous work showed no color-sweetness interactions when unflavored solutions of distilled water, food-coloring, and sucrose were evaluated.[2] With pear-nectar, however, there was a pronounced tendency for *O*s to designate green-colored samples as the least sweet when, in fact, all samples contained 5% sucrose.

The present investigation was undertaken to determine whether the same interaction of color and sweetness could be demonstrated with an entirely different panel of trained *O*s. In addition, the experiment was extended to include blue coloring (Experiment I) and to test for sourness-color relations (Experiment II).

Selection and training of observers. The panel was composed of 11 *O*s selected on the basis of high acuity in recognizing differences in sweetness established in previous testing. Training sessions on uncolored pear-nectar were held daily at 3 P.M. for 5 days. All judging was done in individual, temperature-controlled, partitioned booths illuminated with 150-w. 'day-light' bulbs. Each day, *O*s received 4 randomly presented sets and were requested to circle the number of the sample within each pair that was sweeter (Experiment I) or more sour (Experiment II).

Preparation of samples. Experimental samples of canned, unsweetened pear-nectar (45% Bartlett pears and 55% water) were used throughout the study. This nectar had a content of natural soluble solids of 6.47° Brix as determined by refractometry, and a total acidity of 0.136% calculated as citric acid anhydride. In Experiment I, uncolored. red, yellow, green, and blue nectars (0.8 ml. commercial food-coloring/100 gm. nectar) at 5.0 *vs.* 5.2% sucrose were compared for sweetness. The 25 paired combinations were evaluated in sets of 4 using a balanced, incomplete-block design described by Cochran and Cox.[3]

Five days of testing were required to complete the design. On the following 5

---

* Received for publication August 25, 1961. This investigation was supported in part by a grant-in-aid by the Sugar Research Foundation, Inc., New York, N.Y.

[1] B. I. Masurovsky, The right food color, *Food Eng.*, 11, 1939, 13; H. G. Schutz, Color in relation to food preference, in K. T. Farrel, J. R. Wagner, M. S. Peterson, and G. M. MacKinney (eds.), *Color in Food*, 1954, 21.

[2] R. M. Pangborn, Influence of color on discrimination of sweetness, this JOURNAL, 73, 1960, 229-238.

[3] W. G. Cochran and G. M. Cox, *Experimental Designs*, 1957, 543.

days, samples identical in sugar content (5.0% sucrose) but differing in color were evaluated by the panel. The *O*s were informed that the difference between samples was 5.0 *vs.* 5.1% sucrose, so as not to destroy their initiative for trying to detect differences.

Experiment II was designed to establish the effect of color on discrimination of sourness in pear-nectar, using the same pattern described above. The sweetness-level was 5.0% sucrose for all samples, and sourness was adjusted by additions of 0.01% citric acid. Seven sessions had been used to train the *O*s to detect small differences in sourness and to establish accurancy of discrimination in uncolored samples. Five days of testing completed the experimental design, after which samples identical in acid-content but differing in color were presented as described for Experiment I.

Results: (1) Taste-discrimination in colored and uncolored solutions. Table I lists the panel's reactions to sweetness and sourness in colored and uncolored pear-nectar. Without exception, correct identification of the sweeter sample within pairs was more frequent in uncolored than colored nectars. Results were similar in evaluation of acidified samples except for two *O*s. The total difference between colored and uncolored nectars in percentage of correct responses was significant at the 0.1% level and at the 5% level for the sucrose- and acid-series, respectively. Thus the introduction of visual differences appears to have reduced the accuracy of response to gustatory stimuli. Men, as a group, had slightly more correct responses than did the women. The grand average indicates that the differences in taste between 5.0 and 5.2% sucrose and between 0.0 and 0.1% acid are very subtle. These small differences were considered desirable by the investigator to emphasize the effect of color on taste-discrimination.

TABLE I

RESPONSE TO SWEETNESS AND SOURNESS IN COLORED AND UNCOLORED NECTAR
(Values represent percentages of correct responses)

Stimuli	Unicolored nectar‡	Colored nectar	*t*
5.0 vs. 5.2% sucrose	65.4± 4.9	50.0±11.8	3.971†
0.0 vs. 0.01% acid	80.6±16.9	65.9±10.2	2.469*

* Significant at the 1% level.
† Significant at the 0.1% level.
‡ Training panel.

(2) Color-taste associations. The *O*s were differentially influenced by specific colors, as detailed in Table II. No consistent pattern of response emerged between the studies of sugar and of acid, between the men and the women, or even within the same *O* from one set to another. There was a slight tendency for one woman to associate sourness, and for another to associate sweetness, with the blue-colored samples. Less sweetness was ascribed to red-colored nectars in Set I but not in Set II, where less sweetness was associated with blue and colorless samples. There was no tend-

ency to designate green-colored nectar as less sweet or colorless nectar as more sweet, as was the tendency with a previous panel.[4] Color-taste associations, if manifested at all in a product such as pear-nectar, apparently differ considerably among individual *O*s.

(3) Order of presentation of samples. Midway through the study in which sucrose was a variable, it was observed that certain *O*s were exhibiting a positional bias, *i.e.* excessive selection of the first or second sample within a pair, despite complete randomization of the order of presentation. Consequently, a record was kept of the frequency with which the first member of the pair was selected as having greater intensity of taste. The positional bias differed considerably from one *O* to another. In general, a negative time-error (bias for the second sample) was demonstrated for the group. These observations differ from those reported by Mitchell, who found a significant positive time-error, the magnitude of which was directly related to the amount of difference in taste between the paired samples.[5]

TABLE II

EFFECT OF COLOR ON RESPONSE TO SWEETNESS AND SOURNESS
(Values represent distribution of responses expressed in percentages)

	Men	Women	Total
Set I: 5.0 *vs* 5.2% sucrose			
Colorless	17.8	23.8	21.0
Red	13.5†	18.8	16.0
Yellow	22.9	16.2	20.0
Green	22.9	20.0	23.5
Blue	22.9	21.2	19.5
Set II: 5.0 *vs* 5.0% sucrose			
Colorless	17.2	16.2	16.0
Red	22.1	18.8	21.0
Yellow	22.9	16.2	21.0
Green	22.1	25.0	23.5
Blue	15.7	23.8	18.5
Set III: 0 *vs* 0.01% acid			
Colorless	19.3	30.0*	23.0
Red	22.9	18.8	20.0
Yellow	18.6	13.8	18.0
Green	22.1	15.0	20.0
Blue	17.1	22.4	19.0
Set IV: 0.01 *vs* 0.01% acid			
Colorless	20.0	21.2	19.0
Red	17.9	12.5†	18.0
Yellow	20.7	20.0	20.0
Green	22.1	18.8	20.5
Blue	19.3	27.5†	22.5

* Significant at the 1% level.
† Significant at the 5% level.

---

[4] Pangborn, *op. cit.*, 237.

[5] J. W. Mitchell, Time-errors in the paired comparison taste preference test, *Food Technol.*, 10, 1956, 218.

PERFORMANCE OF NEONATAL MONKEYS IN THE VISUAL CLIFF-SITUATION

By LEONARD A. ROSENBLUM, State University of New York,
and
HENRY A. CROSS, Oklahoma State University

Presented with the opportunity of stepping onto a transparent surface over a shallow area as opposed to one which is suspended a considerable distance off the floor, young animals of a variety of species choose the shallow side.[1] Their work on this problem has led Gibson and Walk to the conclusion that the survival of a species requires the development of depth-discrimination by the time independent locomotion appears. In the rhesus monkey, the early presence of rudimentary locomotor ability and a partially arboreal habitat suggest that high survival-value would attach to the early appearance of visual depth-discrimination. The present study was designed to measure the appearance of differential reaction to depth in the neonatal rhesus monkey during the first 20 days of life.

Subjects. The *Ss* were 11 infant rhesus monkeys born at the Wisconsin Regional Primate Research Center. Each *S* was separated from its mother during the first two days post-partum and housed individually in a wire-mesh cage whose dimensions were 24 × 19 × 15 in. A heating pad was placed on the floor of each cage during the first 15 days of the experiment.

Apparatus. The test-situation was a wooden chamber, 53 in. high, 46 in. wide and 10 in. deep. The chamber was divided vertically into halves by an opaque partition extending from the floor to within 9 in. of the top, and at this height the entire apparatus was covered by a clear Plexiglas shelf. Two identical masonite panels, 23 in. long and 10 in. wide, were painted with a checkerboard of green and white squares, each 1 in. on a side. One panel was placed flush on the underside of the Plexiglas shelf and the other placed on the floor of the apparatus, 44 in. beneath the Plexiglas surface. The entire apparatus was illuminated by a twin 36-in. fluorescent reflector fixed 15 in. above the center of the Plexiglas; thus, during testing there was a single Plexiglas surface which presented a floor directly beneath its surface for half its extent and an apparent drop on the other half, the so-called 'visual cliff' side. During the choice-trials, a 10-in. sq. plywood panel covered with black, coarse sandpaper was centered on the Plexiglas surface. Latencies on choice-

---

* Received for publication August 24, 1961. This research was done at the Wisconsin Regional Primate Research Center. It was supported by Grants M-772 and RG-4857 from the National Institutes of Health, and by the Graduate School of the University of Wisconsin.

[1] E. J. Gibson and R. D. Walk, The visual cliff, *Sci. Amer.*, 202, 1960, 64-71.

trials, defined as the time elapsed from the placement of *S* on the sandpaper until a choice was made, were obtained by means of a Standard Electric timer, operated by a hand-switch held by *E*. A synchronous-motor timing device connected to a small bulb unseen by *S* indicated 15-sec. intervals to *E*.

Procedure. Every *S* was tested daily throughout Days 3-20 of life. Each day's testing consisted of two direct-placement trials separated by two choice-trials, with the order of the trials alternated from day to day for each *S*. A direct-placement trial consisted of the following: *S* was placed directly onto either the visual cliff or the shallow side of the apparatus, its head facing the opposite side, and allowed to remain for 90 sec. The right-left position of the shallow side and visual cliff side was alternated daily. Each trial was divided into six 15-sec. periods, and *E* recorded on a check-sheet the appareance or nonappearance of a series of behaviors previously shown to be indicative of emotional disturbance in infant rhesus monkeys.[2] The four behaviors recorded in this context were crouching, vocalization, self-clasping, and rocking (rhythmic, stereotyped movements of the body or its parts). Three disturbance-levels were recorded, a zero-level, a low level, and a high level, characterized by no disturbance, a single disturbance, and two or more disturbances, respectively. These levels were given numerical values of 0, 1, and 2. A second kind of measure recorded during each 15-sec period was 'freezing,' defined as a failure to locomote or to change position.

The procedure for the choice-trials was as follows: The sandpaper-covered panel was positioned; *S* was placed on it, with its head parallel to the two sides, and then allowed 200 sec. to make a choice, each choice being defined as the movement of any three limbs off the sandpaper. When a choice was made, the trial was terminated. Intertrial intervals were approximately 2 min.

The significance of differences for all tests was measured by analysis of variance. Since the data did not meet all the assumptions underlying the *F*-test, a confidence-level of 0.01 was set, in keeping with the general recommendations of Lindquist.[3]

Results. Table I shows the mean disturbance-scores for the two direct-placement trials. The analysis of variance indicated a significantly higher frequency of disturbance on the visual cliff than on the shallow side. The failure of the interaction to reach significance showed that there was no significant change in this differentiation during the 18 days of testing. Similar results were obtained for freezing (Table I), the difference again significant and the interaction not.

For the choice-trials, the percentage of responses made to each side, and the percentage of trials on which there was failure of choice, are shown in Table 1. The difference between visual cliff and the shallow side was significant while the interaction was not. (A number of accidental and presumably spurious responses were made on the choice-trials. These occurred as a result of the *S*'s bolting off the sandpaper without looking, or backing off the sandpaper onto the visual cliff.) One further indication of a clear-cut preference for the shallow side is the fact that of the 11 animals tested, 9 made fewer than three choices of the visual cliff during the 18 days of testing.

---

[2] H. F. Harlow and R. R. Zimmerman, Affectional responses in the infant monkey, *Science*, 130, 1959, 421-432.

[3] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1956, 78-86.

Discussion. Locomotion, climbing, and even jumping appear early in the behavioral repertoire of rhesus monkeys. Since, in the feral state, part of this behavior must occur in trees some distance from the ground, the importance of the ability to discriminate depth for the survival of the young rhesus is obvious. Despite the fact that the first days of life of the infant are spent in comparative safety at the mother's breast, the neonate seems capable of discriminating and reacting to visual depth. Further work is required to analyze the cues on which the discrimination revealed in the present study is based.

TABLE I

MEAN RESPONSE IN THE VISUAL CLIFF-SITUATION AS A FUNCTION OF AGE

	Age in days					
	3–5	6–8	9–11	12–14	15–17	18–20
Disturbance						
cliff	1.40	1.50	1.55	1.42	1.41	1.37
shallow	1.16	1.16	1.20	1.26	1.09	1.13
Freezing						
cliff	2.83	3.04	3.22	3.11	2.85	3.15
shallow	2.32	1.61	1.96	2.00	2.04	1.71
% choice						
cliff	11.3	3.3	9.3	9.0	3.6	0.0
shallow	62.0	65.0	55.6	59.3	59.3	57.3
No response	26.7	31.7	35.1	31.7	37.1	42.7

SUMMARY

Eleven infant rhesus monkeys were tested on an adaptation of the 'visual cliff' in the third through the twentieth days of life. Discrimination of depth was evident on the third day and showed no significant change thereafter.

APPARATUS

AN APPARATUS FOR STUDIES OF BRIGHTNESS WHICH CONTROLS VARIATIONS OF INTENSITY WITH TIME

By PAUL R. ESKILDSEN, University of Oregon

It has been shown that the spacing and frequency of presentation of discrete stimuli in categorical rating-scales has an influence on the resulting psychophysical function.[1] Stevens describes an iterative technique giving a 'pure' or 'unbiased' scale where stimuli are distributed in such a way that the categories are equally spaced along the psychological continuum. For continuous stimuli there may be an analogous 'biasing' effect upon the psychophysical relationship.[2] In psychophysical techniques where a continuously variable stimulus is required, as in the method of adjustment, there is rarely any experimental control over the variation of the physical stimulus with time. In the case of brightness, presumably the minimal bias would result when the stimulus was varied in such a way that $R = at$, where R is the psychological magnitude (brightness) of the stimulus, and t is time.

The apparatus described in this report was designed as a general psychophysical device for investigating brightness. Specifically it can be used: (a) for experiments studying the effects on brightness-judgments of variations in stimulus-intensity with time; (b) for brightness-scaling using psychophysical procedures requiring continuously variable stimuli such as in equisection and prescribed-ratio procedures;[3] and (c) for presentation of single or multiple discrete stimuli permitting the use of procedures of direct magnitude estimation; and (d) the method of constant stimuli for measurement of *jnd*s.

Fig. 1 is a schematic diagram of the lateral view of the apparatus. Illumination is provided by a single source, A, a 500-w., Sylvania projection-lamp with regulated voltage-input. The source is placed in front of a curved reflecting panel, B, which forms the rear of a light-chamber, the interior of which is painted with Iverson's

---

* This research was supported by a grant from the National Science Foundation under the direction of D. Robert Fagot. The author is indebted to Dr. Raymond Hohle for his assistance in the design and construction of the apparatus.

[1] S. S. Stevens and E. H. Galanter, Ratio scales and category scales for a dozen perceptual continua, *J. exp. Psychol.*, 6, 1957, 377-411.

[2] W. R. Garner, Context effects and the validity of loudness scales, *ibid.*, 48, 1954, 218-224; S. S. Stevens and E. C. Poulton, The estimation of loudness by unpracticed observers, *ibid.*, 51, 1956, 71-78.

[3] W. S. Torgerson, *Theory and Methods of Scaling*, 1958, 94-95, 118.

XP 354, a specially mixed, heat-resistant, high-reflectance (98%), flat white paint. Four triangular apertures are cut into the front panel of the source-chamber such that diffuse illumination is seen over the entire area of each of the triangles. One of these apertures is indicated in the figure at C. Shutters, actuated by levers, D, from the rear of the apparatus operate to cover or uncover the apertures, making progressively smaller or larger triangles, thus permitting more or less light to pass through. The light which does pass through falls on opal-glass diffusing screens, E, which are the display-targets seen by *S*. The four diffusing screens are separated by light-tight barriers extending to the front panel of the source-chamber. Direct light from the source does not fall on the opal-glass screens. At the rear of the apparatus the levers which control the amount of light emitted have pointers attached for reading level of intensity from the calibrated scale, F.

It was desired to have sensitivity in control of the intensity of light such that possible errors in readings on the calibrated scale or errors in setting of the intensity would be more nearly equal all along the psychological continuum than would be the case if the scale were linear with the physical continuum. For this reason a triangular aperture was used in combination with a slight displacement of the source which results in more light being reflected through the lower portion of the triangles than the upper. This arrangement gives a calibrated scale which varies as the intensity to the 0.275 power.

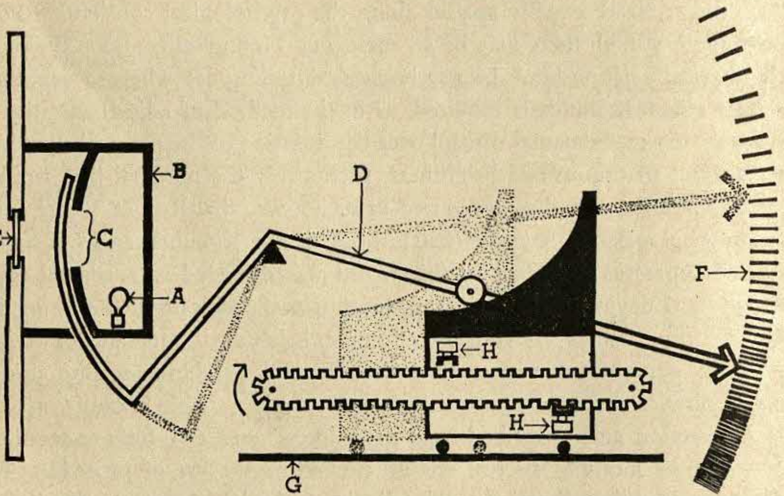


FIG. 1. SCHEMATIC DIAGRAM OF APPARATUS FOR SCALING BRIGHTNESS SHOWING OPERATION OF CAM-DEVICE

The levers which control the intensity can be raised and lowered either manually by *E* or by means of a specially constructed cam-device. This device consists of a track (G in Fig. 1) upon which slides a movable carriage which has mounts for holding the cam in position. Alongside the carriage runs a roller-chain driven by a constant-speed eleceric motor at the rate of 42 cm. per min. Two solenoids, *H*, on the carriage are equipped with grippers which grasp either the forward-moving or the backward-moving portion of the roller-chain as desired, pulling the cam along.

The shape of the cam then determines how the lever is raised and lowered and thus the functional relationship which intensity has with time. By changing cams, the increase in intensity with time can be made to match any assumed psychophysical function. Both *S* and *E* are provided with a control-switch for operating the cam-device. The length of the track permits up to 45 sec. total time for the cam to change the brightness from one extreme to the other. The entire mechanism, source-chamber, levers and cam-device are enclosed in a cabinet 25 × 30 × 48 in. in size.

S is presented with four luminous sources, each of which has continuously variable illumination, with the color-temperature of each source equal and constant throughout the entire range of illumination.[4] By means of a switch, *S* can control the intensity of one of the sources, making it brighter or dimmer as he desires. The four sources are circular, 13 mm. in diameter, and placed horizontally with a 90-mm. separation. The surround is painted flat black. A display-shutter makes possible the presentation of any number and combination of the sources. Intensity can be varied continuously from 0.05 to 580 ft.-lam. For most experiments this range is sufficient to make it unnecessary to use neutral-density filters to extend the range. Five readings were taken with a MacBeth Illuminometer on each of 23 points along the scale for each of the four sources. The results were averaged and plotted and smooth calibration-curves drawn through the points.

The cams are cut to provide a uniform change in brightness of the light according to various assumed psychological scales, such that $R = at$, where *R* is the psychological magnitude of the stimulus and *t* is time. Following Stevens' assumption that the psychophysical law for brightness is a power-function, then $R = bS^k + c$ where *S* is the physical magnitude of the stimulus and b, c, and k are parameters peculiar to the stimulus continuum and the individual being tested.[5] What is desired is a stimulus which will vary such that $S = (at - c)^{1/k}/b$, thus giving a uniform change in brightness with time. For the more general case where the stimulus-range has the minimal and maximal psychological limits of R_L and R_H and where the total time necessary to go from R_L to R_H is $t_{max.}$, then $(R_i - R_L)/(R_H - R_L) = (t_i/t_{max.})$ and $R_i = (t_i/t_{max.})\ (R_H - R_L) + R_L$. Converting to physical values, $S_i = [(t_i/_{max.})\ \ (S_H^k - S_L^k) + S_L^k]^{1/k}$. For any particular range of stimuli, $S_H - S_L$, and any particular rate determined by $t_{max.}$, a series of cams can be cut corresponding to different assumed exponents of the psychological power-law. For a logarithmic law the function is *log* $S_i = (t_i/t_{max.})\ (\log S_H - \log S_L) + \log S_L$.

In controlling the function by which the stimulus is increased or decreased a new psychophysical procedure is made possible in which direct judgments can be obtained from *Ss* of the uniformity of the rate of change in brightness. For example, a paired-comparisons procedure could be employed using a series of cams to determine the exponent giving a linear change in brightness. In many instances apparatus is constructed for use in only one experiment, and must be modified when experimental procedures are changed. It is hoped that the device described here will be versatile enough to be useful for many different experimental procedures without need for redesigning.

---

[4] A previous apparatus, constructed with four different 100-w. sources for each of the four targets, proved unsatisfactory because of slight differences in color-temperature between the bulbs, which affected the psychophysical scale. Different rates of aging of the four bulbs also resulted in unsatisfactory results after only a few hours of use, making a single source for all targets desirable.

[5] Stevens. On the psychophysical law, *Psychol. Rev.*, 64, 1957, 153-181.

APPARATUS NOTES

THE TAPE-RECORDER AS AN INSTRUMENT FOR MEASURING TIME

We have found, by methods devised in our laboratory, that the tape recorder is a useful instrument for measuring reaction-times and for timing success and failure in studies with the pursuit-rotor.[1]

Reaction-times. Reaction-times may be obtained as follows. The primary winding of a 110 to 6.3-v. filament transformer is energized simultaneously with the onset or the visual or auditory stimulus. This is accomplished by wiring the primary in shunt with the visual or auditory stimulus or by means of wiring the primary winding to one side of a double pole single throw switch while the visual or auditory stimulus is wired to the second side. The secondary winding, in conjunction with a full wave-rectifier supplies a 120 pulse per second signal to the 'phono input' jack of the tape-recorder. *S* reacts to the stimulus by operating a switch that breaks the primary circuit and thus ends the recorded signal.

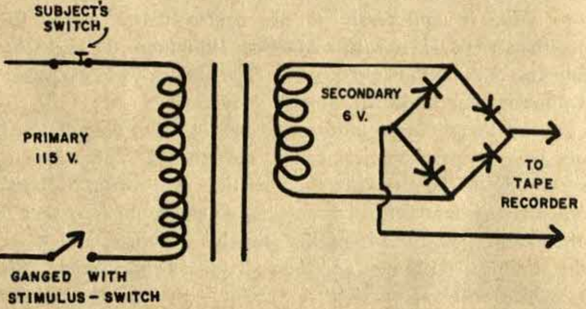


FIG. 1. STEP DOWN TRANSFORMER WITH RECTIFIERS
Parts: 110v.-6.3 Filament Transformer. Four IN, 34 Diodes.

Following the experimental session, the tape is run through the circuit shown in Fig. 2 which is connected to a stop-clock.[2]

When the tape is run through this circuit, Relay 1 closes and the clock starts. With its termination, Relay 1 opens and the clock stops, giving the measurement of *S*'s reaction-time.

The circuit is conventional except for the large value of R_4. This value is required to bias the pentode to a point where Relay 1 will not operate on no-signal input. R_5 has been made adjustable to compensate for differences in relays that may be used. C_2 prevents degenerative feedback over R_4 making the device quite sensitive. C_3 prevents the relay from chattering on 60-v. input signals. The chassis of the device is not used as common or ground for the circuit thus the shock hazard is negligible. The mechanical delay of the relay has been electronically measured and found to be 0.0027 sec.

The accuracy of the method has been determined by simultaneously recording a series of reaction-times on the tape recorder and measuring them directly by the clock. Comparison of the measurements obtained directly by the clock and those obtained by the tape recorder method have indicated, by the complete agreement in readings, that the tape recorder method will measure reaction-times accurate to 0.01 sec.

---

* This work was supported by a grant-in-aid by the Wallace Laboratories, New Jersey.

[1] Philips Tape Recorder, Model 400.

[2] Lafayette, No. 202254.

To insure accuracy at this level, tape recorders must be uesd whose tape transport speed variations are 0.3% or less. It is also essential that the tape used have a yield strength that is close to their tensile strength so that permanent stretch deformation produced by jerking the tape or by temporary defect in the recorder will be signalled by the tape's breaking. For this purpose, tapes with a cellulose acetate backing rather than a polyester backing are necessary.

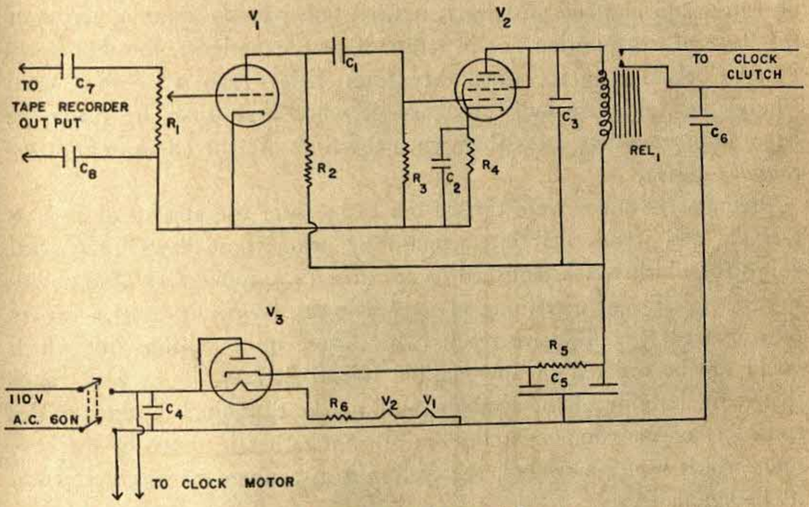


FIG. 2. WIRING DIAGRAM TO CLOCK

Parts $V_1 = 12$ AT6; $V_2 = 50$C 5; $V_3 = 35$W4; Rel.$_1 = 5000$ Ohm relay (Potter and Brumfield L.B.5); $C_1 = 0.1$ MFD, 200 v.; $C_2 = 8$ MFD, 50 v.; $C_3 = 1$ MFD, 200 v.; $C_4 = 0.05$ MFD, 200 v.; $C_5 = 40\text{-}40$ MFD, 150 v.; C_6, C_7, and $C_8 = 0.1$ MFD, 200 v.; $R_1 = 500$K potentiometer; $R_2 = 300$K ¼w.; $R_3 = 500$K ¼w.; $R_4 = 10$K. Wire wound potentiometer $R_5 = 1100$ Ohm 2 W.; $R_6 = 125$ Ohm 10 W.

The advantages of this method of measuring reaction-times are: (1) The silence of the method during recording on tape. The noise of clocks starting often produces serious problems in laboratories. (2) With little practice, the tape-recorder can be operated in complete darkness. (3) *S*'s reaction-times are retained on tape for as long as the *E* requires.

Pursuit-rotor. Other timing problems are solved by the use of this procedure. In studies with the pursuit-rotor, for example, most *Es* are usually content with recording the percentage of the time that *S* is on the target during the periods of the experiment. The measurement of the number of hits as well as percentage time-on-target involves elaborate apparatus. By the use of the tape recorder, however, one may have percentage time-on-target, number of hits, and length of time-on-target during each hit. Here also, the silence of the method is a big advantage.

C. G. COSTELLO
JAMES STEPHEN

Regina General Hospital
Saskatchewan, Canada

NOTES AND DISCUSSIONS

GRAPHIC SYMBOLS: THINGS OR HOLES?

This note reports a set of observations on the figural characteristics of letters and numbers. Contrary to what one would expect in terms of the 'laws of organization,' or in terms of past experience, the letters appear to be easier to see when they form 'holes' than when they form 'things' (*i.e.* they are more recognizable when their lines are made the edges of *surrounding* areas than when the lines are the edges of the *surrounded* areas).

Preliminary studies were carried out to measure the efficacy of various *Gestalt* 'laws' (and their interactions) by using those 'laws' to conceal forms for which well-established names exist (*e.g.* numerals), thus avoiding the necessity of a training period.[1] The *O*s, trying to detect a variety of concealed symbols, gave results which were quite reliable, but which could not be ordered in any sensible way. Belatedly, it occurred to us that recognizability of the symbols could not be explained in terms of detectability of the component lines at all, but rather in terms of the *contours* which *bound* a symbol, segregating it as a phenomenal object from its background.[2]

The first two experiments (Experiments I and II), were designed to test the hypothesis that an 'open' 4 (Fig. 1, Sets 1a and 2a) was concealed in various ways; the lines of the top portion of the 4 being made into edges of the surrounding areas which are 'outside' the body of the 4 (Patterns b and c of Sets 1 and 2). Conversely, in Patterns d, e, and f of these sets, the same lines were made into edges of surrounded 'inside' areas.

Each of the two groups of *S*s (48 in all) was presented with one set of variations and given the following instructions:

> The figure 4 appears at the extreme left; it is also embedded in the rest of the figures. Place the letter corresponding to the figure in which the 4 is *least* easy to see. at *O* on the scale below; place the most visible 4 at 10 on the scale, and the labels of the rest where appropriate between these extremes.

---

[*] This work was supported by Grants NSF G-4458 and G-9601 from the National Science Foundation.

[1] A number of the more-or-less casual *Gestalt* demonstrations depend upon concealing letters and numbers by the use of 'good continuation' (Wolfgang Köhler, *Gestalt Psychology*, 1929, 210-211). The concealment of familiar, named geometrical shapes was studied by A. Galli and A. Zama (Untersuchungen über die Wahrnehmung ebener geometrischer Figuren; die ganz oder teilweise von anderen geometrischen Figuren verdekt sind, *Z. Psychol.*, 123, 1931, 308-348).

[2] Wolfgang Metzger, *Gesetze des Sehens*, 1953, 11 f, 51.

Similar instructions were used in all of the experiments reported here. The results are given in Table I.

In all the experiments, visibility was reported as greater for Patterns b and c, than for d, e, and f; a nonparametric analysis of variance shows the differences between Patterns b through f to be significant at the 1% level in each set.[3]

Is 'good continuation' responsible for these differences? If we try to partial out the 'good continuation' which might conceal the characteristic left-hand free corner of the 4 we note that, free of such concealment, $c > e$ ($p < 0.01$), and (with such concealment in both cases, we still find $b > d$ ($p < 0.01$); moreover, use of 'good continuation' to conceal the

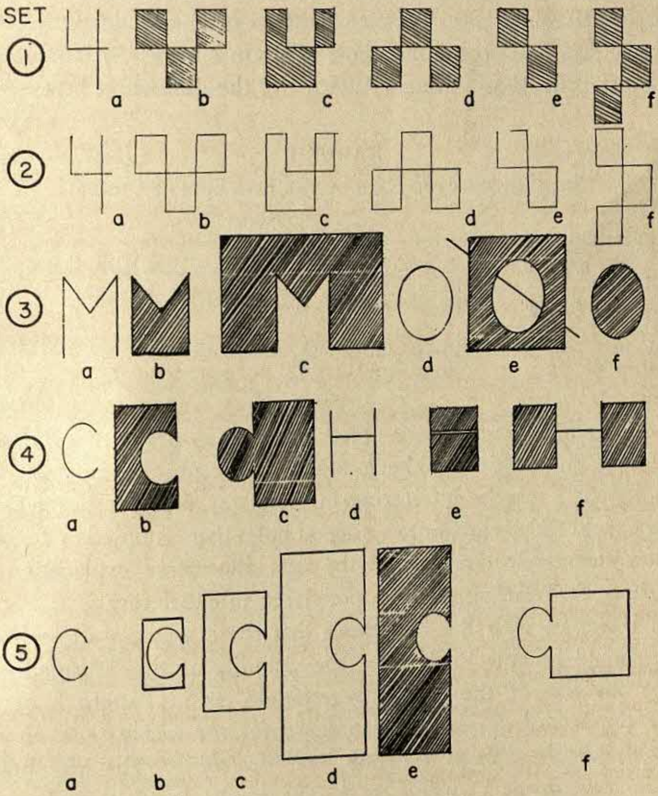


FIG. 1. THE FIVE SETS OF STIMULUS-OBJECTS

tail shows no difference between Patterns e and f. 'Good continuation' does not seem to explain these findings.

---

[3] Milton Friedman, The use of ranks to avoid the assumption of normality, *J. Amer. stat. Assoc.*, 32, 1937, 675-701.

At this point, two alternative explanations must be considered: (1) The open top of the 4 is a critical characteristic for its recognition, which is absent from Patterns d, e, and f in Sets 1 and 2. (2) The lines of the 4 normally act as contours facing inward, *i.e.* the body of the 4 normally acts as a 'hole' rather than a 'thing.'

Experiments III and IV were designed to test these alternatives, using the stimuli shown in Sets 3 and 4, with the results shown in Table I. In Experiment III, differences between Patterns b, c, e, and f in Set 3 were significant, $p < 0.01$; the *M*-as-a-hole (Pattern *c* in Set 3) is more visible ($p < 0.01$) than *M*-as-a-thing (Pattern b); the differences between *O*-as-a-hole (Pattern e) and *O*-as-an-object (Pattern f) are not significant, but the diagonal line (which had been intended to insure the 'groundal' nature of the 'infield') may be interfering in some way. In Experiment IV, (Set 4), the *C*s and the *H*s are more visible as holes than as objects ($p < 0.01$). We now tentatively can discard the first alternative explanation given above, since the opening of *'C'* is not closed in either case.

TABLE I

MEAN RATINGS FOR EACH OF THE FIVE SETS OF STIMULI

Experiment	Stimulus set	Pattern a	b	c	d	e	f	*N*
I	1	10	4	8	2	3	2	23
II	2	10	5	8	2	4	4	25
III	3	9	1	7	9	4	3	30
IV	4	10	8	1	10	0	5	27
V	5	9	8	7	4	3	0	27

This leaves us with the hypothesis that these symbols act as holes. If so, we should be able to decrease the visibility of the 'C' simply by increasing the size of the enclosing rectangle, as in Pattern b of Set 5, *because this would decrease the tendency for the rectangle to be figure,*[4] *and, consequently, change 'C' from 'hole' to 'thing.'* The hypothesis was tested in Experiment V, using Set 5. The results for Patterns b, c, and d (given in Table I) differ from each other ($p < 0.01$).

The letters examined here appear to act as 'holes' rather than as 'things,' and a large-scale investigation of this phenomenon seems to be called for, extending, on the one hand, to all of the letters and numbers in various type-faces, and, on the other, to combinations of letters into words and phrases, since the present findings are both consequential and puzzling. It is important to know that symbols act as 'holes,' if they do so in general, since the requirements of visibility and concealment are quite different (in fact, reversed) for 'holes' and for 'things.'

---

[4] Tadasu Oyama, Determining factors of figure-ground perception, *Hokkaido Rep. Psychol.*, 2, 1960, 1-24.

It is puzzling that letters should act this way, since (a) in terms of the Gestalt 'laws,' we should expect letters, which are small enclosed and almost-closed regions, to be figure (thing) rather than ground (hole); and (b) in terms of past experiences with block letters and the like, and with the necessity for recognizing the same letters as they are shuffled around spatially against the same or different backgrounds, we should expect them to be identified as things-with-edges, rather than as holes-behind-edges.

Cornell University
Julian Hochberg
Virginia Brooks

LOSS AND RECOVERY OF DISCRIMINATION OF VISUAL DEPTH IN DARK-REARED RATS

Stimulation by light has been shown to be essential for the normal development and maintenance of pattern-vision in chimpanzees and kittens.[1] Recovery or partial recovery of lost pattern-vision after re-introduction into light has been reported for chimpanzees;[2] similarly, kittens have shown development of pattern-vision after introduction into light.[3] Such relationships have not, however, been shown for rats. Lashley and Russell showed that rats raised in darkness to 100 days of age discriminated distance with little or no training,[4] and Hebb found that rats raised in darkness "to maturity" discriminated patterns.[5] More recently, Walk, Gibson, and Tighe, and Nealey and Edwards failed to detect differences between 90-day-old light-reared and dark-reared rats tested for discrimination of depth on the 'visual cliff' devised by Walk *et al.*[6]

---

[1] A. H. Riesen, Arrested vision, *Scientific American*, 183, 1950, 16-19; E. J. Gibson and R. D. Walk, The 'visual cliff,' *ibid.*, 202, 1960, 64-71.

[2] Riesen, *op. cit.*, 17, 19.

[3] Gibson and Walk, *op. cit.*, 71.

[4] K. S. Lashley and J. T. Russell, The mechanism of vision. XI. A preliminary test of innate organization, *J. genet. Psychol.*, 45, 1934, 136-144.

[5] D. O. Hebb, The innate organization of visual activity: I. Perception of figures by rats reared in total darkness, *ibid.*, 51, 1937, 101-126.

[6] R. D. Walk, E. J. Gibson, and T. J. Tighe, Behavior of light- and dark-reared rats on a visual cliff, *Science*, 126, 1957, 80 f.; S. M. Nealey and B. J. Edwards, 'Depth perception' in rats without pattern-vision experience, *J. comp. physiol. Psychol.*, 53, 1960, 468 f.

Recently we had an opportunity to compare the depth-discrimination of rats reared in the dark for 300 days with that of their light-reared controls on a 'visual cliff'. Contrary to previous findings with rats, the dark-reared animals failed to discriminate, but the light-reared ones discriminated as usual. The details are as follows.

Procedure. Thirty light-reared and 30 dark-reared 300-day-old rats were tested for depth-discrimination on a visual cliff which was a modification of one described by Walk and Gibson.[7] The basic design of this apparatus was the same as that of other visual cliffs—a runway centered on a plate-glass surface with a 'near' surface directly under the glass on one side, and a 'far' surface farther below the glass on the other side. The apparatus differed from the 53-in. visual cliff originally described by Walk *et al.* in three main ways.[8] First, the present apparatus was designed to eliminate density of pattern as a cue to depth by the use of near and far patterns that subtended equal retinal angles. Secondly, the depth of the far side was reduced from 53 in. to 10 in. lower than that of the near side in order to insure that the pattern on the far side would be visible to the rat. Finally, cues from motion-parallax were minimized by the use of patterns consisting of three lines parallel to the runway on each side of it. Under these conditions cues from motion-parallax are still produced by movements of the head made perpendicularly to the runway, but cues from movement along the runway are eliminated. Illuminance at the rat's eye was the same on both sides of the runway.

Each of the 60 rats was taken individually from the dark and given 20 min. to adapt to the light before Test 1. Following light-adaptation, each rat was placed on the runway and observed for 5 min., or until it stepped down on the near or far side. Following its test each dark-reared animal was immediately returned to the dark.

Results. The results of this test are shown as Test 1 in Table I. The proportion of light-reared rats selecting the near side is reliably greater than the proportion selecting the far side ($\chi^2 = 8.00$, $p < 0.01$); the performance of the dark-reared rats is at chance. For reasons that we do not understand and have not explored, the number of animals failing to respond was greater in this test than in previous tests with the 53 in.-cliff of Walk *et al.*[9]

To determine whether the performance of the rats reared in darkness was due to their rearing or to a reduction in cues to depth in this apparatus, we tested the dark-reared rats three days later with Nealey and Edwards' copy of the original 53-in. visual cliff of Walk *et al.*[10] The results of this test, Test 2 in Table I, show that the animals again performed at a chance level. This finding is to be compared with Nealey and Edwards' earlier finding that 80% of their 90-day-old dark-reared rats chose the near side. The same apparatus and the same strain of rats were used in both studies.

---

[7] Walk and Gibson, A comparative and analytical study of visual depth perception, *Psychol. Monogr.,* 75, 1961, (No. 519), 1-44.

[8] Walk, Gibson, and Tighe, *op. cit.,* 80 f.

[9] Walk, Gibson, and Tighe, *op. cit.,* 80 f.

[10] Nealey and Edwards, *op. cit.,* 468.

Further observations. Following Test 2, 14 rats from each group were selected randomly, given atropine, and subjected to an ophthalmoscopic examination. No differences in the retinae of the dark- and light-reared groups were observed.

The remaining 16 dark-reared as well as the 15 light-reared rats were kept in the light for a month after Test 2 and tested again with the 53-in. cliff. The dark-reared rats were then returned to the dark for 18 days, retested and then left in the light, and two further tests with the same apparatus were made, one 42 days after last removal from darkness and another 13 days later. These results are shown as Tests 3, 4, 5, and 6 in Table I. The dark-reared rats appear to show improvement in discrimination with exposure to light (Tests 3, 5, and 6) and to deteriorate in discrimination with a return to the dark (Test 4). The light-reared rats, on the other hand show a deterioration in discrimination with repeated testing (Tests 5 and 6). We have not tested the significance of the differences after the first set of observations because we did not contemplate keeping individual records on our dark-reared animals until after the second test, and by then we thought it was too late. Consequently we have no way of assessing the correlation between successive tests. Nevertheless,

TABLE I

BEHAVIOR ON TWO "VISUAL CLIFFS"

Test	Condition of Testing	*N*	Frequency of choice		
			near	far	no response
			(Visual angles equated)		
1.	D-300 days in dark	30	6	7	17
	L-300 days in light	30	15	3	12
			(Original 53-in. visual cliff)		
2.	D-303 days in dark	30	17	13	0
	L-	—	—	—	—
3.	D-34 days in light since Test 2	16	12	3	1
	L-37 days in light since Test 1	16	15	1	0
4.	D-18 days in dark since Test 3	16	7	9	0
	L-18 days in light since Test 3	16	12	4	0
5.	D-42 days in light since Test 4	16	9	6	1
	L-42 days in light since Test 4	16	9	7	0
6.	D-13 days in light since Test 5	16	11	3	2
	L-13 days in light since Test 5	15	8	7	0

it seems safe to conclude that: (a) dark-reared rats, like cats and apes, show a deterioration of discrimination as a function of time in the dark, since our 300-day-old dark-reared rats did not discriminate depth, but 90-day-old dark-reared rats do discriminate; (b) this deterioration disappears in a cage-environment with normal lights; and (c) continued discrimination depends on continued stimulation by light.

University of California, Berkeley
STANLEY M. NEALEY
DONALD A. RILEY

RESISTANCE TO EXTINCTION AS A FUNCTION OF NUMBER OF REINFORCEMENTS

North and Stimmel, working with rats in a runway, recently reported greater resistance to extinction after 45 than after 135 training trials.[1] By contrast, the bar-pressing data of Williams and Perin yield a function that is asymptotic at about 90 trials.[2] It is of interest to inquire into the effect on resistance to extinction of extended training in bar-pressing.[3]

Eighty experimentally naïve male albino rats of the Sprague-Dawley strain, 69–83 days old at the start of the experiment, were trained in a Davis apparatus (Box DB-103, Lever PD-104). They were assigned at random to one of four experimental groups, Groups A, B, C, and D, which received 10, 90, 180, or 360 reinforced training trials, respectively. All the *S*s were adapted to a 23-hr. food-deprivation schedule for seven days prior to the start of training. On Day 7, each *S* was given 10 45-mg. Noyes pellets in the home-cage prior to being fed its regular ration. Training Day 1 consisted of cup-training during which each *S* received a total of 60 pellets delivered in sessions of 10 pellets each. Between successive 10-pellet sessions, *S* spent 5 min. in his home-cage. On Training Day 2, each *S* received 30 pellets in the same manner, after which it was returned to the home-cage for 10 min. *S* was then brought back to the apparatus which now contained the manipulandum and was allowed to press the bar to obtain an appropriate number of reinforcements—10 times for Group A, and 30 times for Groups B, C, and D. The latter groups continued to receive 30 reinforced bar-presses per day until each had accumulated the appropriate number of reinforcements (90, 180, 360) set for them in the experimental design. One day after the completion of the training, every *S* was extinguished to a criterion of 5 min. without a response.

---

[1] A. J. North and D. T. Stimmel, Extinction of an instrumental response following a large number of reinforcements, *Psychol. Rep.*, 6, 1960, 227-234.

[2] S. B. Williams, Resistance to extinction as a function of number of reinforcements, *J. exp. Psychol.*, 23, 1938, 506-521; C. T. Perin, Behavior potentiality as a joint function at the amount of training and the degree of hunger at the time of extinction, *ibid.*, 30, 1942, 93-113.

[3] This research was supported by Grant M-3331 from the National Institute of Mental Health.

Analysis of variance applied to the data for time required to make the first 10 reinforcement responses showed no differences among the groups. The mean numbers of bar-presses to extinction for Groups A, B, C, and D were, respectively, 104.5, 159.1, 194.1, and 221.3; the corresponding medians were 77.0, 153.3, 173.0, and 220.5. The mean times (in min.) were 47.4, 52.3, 52.5, and 62.4, respectively.

The clear implication of these results is that resistance to extinction of a bar-pressing response is an increasing negatively accelerated function of number of acquisition-trials. Similar results have recently been reported by Harris and Nygaard.[4]

Texas Christian University JAMES A. DYAL
Southern Methodist University THOMAS A. HOLLAND

ON LOTSOF'S PROPOSED RELATION BETWEEN CERTAINTY AND EXPECTANCY

Lotsof has contended recently that "the verbal response of certainty may be employed as a measure of expectancy of success."[1] The contention was based on the results of an experiment in which *S*s were required to rate, on a scale of 1-4, how sure they were of the correctness of each response in a task in which the over-all percentage of correct responses was arbitrarily manipulated. These percentages ranged from 16.6-83.3; one of nine different reinforcement-ratios was used with each of nine different groups. Lotsof's analysis of the data showed a correlation of 0.93 between reinforcement-ratio and mean certainty of response. What his analysis did not show was that the mean certainty-judgments, when converted to percentages, were very nearly the same as the associated percentage of reinforcement values.

This fact is indicated in Fig. 1, where a certainty of 1 is represented as 100% and a certainty of 4 as 0%. With these two points as anchors for a linear transformation, Lotsof's obtained mean certainties all have been converted to percentages and plotted against the reinforcement-ratios. The resultant coördinate points have been fitted by a least-squares line (solid) in Fig. 1. This line is similar to the one (dashed) which would have resulted if the percentage of reinforcement and the percentage of certainty were identical at each point. Thus, not only is there a correlation between the two measures, but the expression of certainty, when converted to percentage of certainty, is a fairly good estimate of the actual percentage of reinforcement. It appears, however, that the experimental conditions were such that the higher percentages of reinforcement generated percentages of certainty that were slightly too low. It also can be argued from the extrapolations of the line obtained by least-squares that a mean judgment of 1 would not represent 'complete certainty' as adequately as a mean judgment of 4 would represent 'complete uncertainty.'

---

[4] Phyllis Harris and J. E. Nygaard, Resistance to extinction and number of reinforcements, *Psychol. Rep.*, 8, 1961, 233-234.

[1] E. J. Lotsof, Expectancy for success and certainty of response, this JOURNAL, 72, 1959, 600-602.

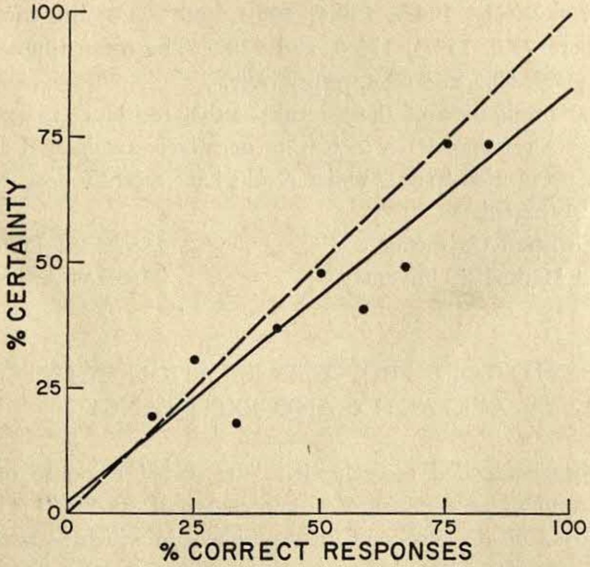


FIG. 1. PERCENTAGE OF CERTAINTY AS A FUNCTION OF PERCENTAGE OF REINFORCEMENT

Indiana University RICHARD N. BERRY

THE RELATION BETWEEN PSYCHOLOGICAL AND MATHEMATICAL PROBABILITY

'Psychological probability' may be defined as perceived mathematical probability. Attempts have been made to measure it in terms of the expected-value model,[1] but, as Edwards has noted,[2] several dubious assumptions are involved. A method of direct measurement is described here, together with some sample results.

The *S*s were given a mimeographed set of 33 problems in probability and asked for a written *estimate* of the mathematical probabilities. The 'correct' answers ranged from 0–1. The *S*s were 41 men and women drawn from a class in introductory psychology. The data of nine of the *S*s had to be discarded because they failed to follow instructions.

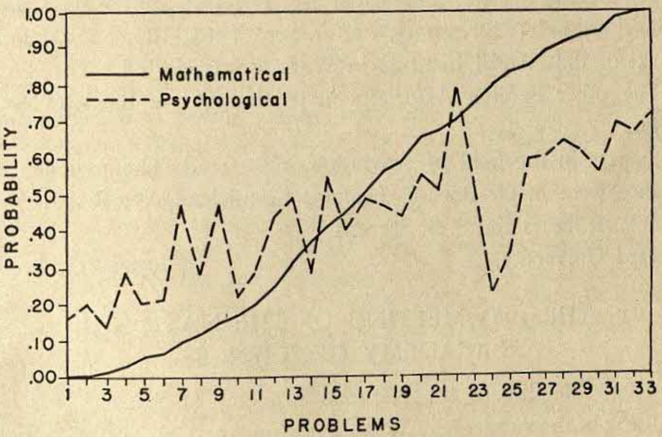


FIG. 1. THE RELATION BETWEEN PSYCHOLOGICAL AND MATHEMATICAL PROBABILITY

The results are presented in Fig. 1. The familiar tendency to overestimate low probabilities and to underestimate high probabilities appears clearly. The point of indifference—in the neighborhood of 0.45—is higher than that reported by most previous investigators.

University of Tulsa THOMAS C. HOWARD

---

[1] R. M. Griffith, Odds adjustment by American horse-race bettors, this JOURNAL, 62, 1949, 290-294; M. G. Preston and Philip Baratta, An experimental study of the auction-value of an uncertain outcome, this JOURNAL, 61, 1948, 183-193.

[2] Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953. 349-364.

THE 1963 MEETING OF THE AMERICAN PHILOSOPHICAL SOCIETY

The Annual General Meeting of the American Philosophical Society was held in Philosophical Hall in Independence Square, Philadelphia, on April 18-20, 1963.

The Karl Spencer Lashley Award was made to Alexander Forbes, Professor of Physiology, Emeritus, of Harvard University in recognition of his useful and significant work in the field of neurobiology. The John F. Lewis Prize was awarded to E. B. Wilson, Professor of Vital Statistics, Emeritus, of Harvard University for an important paper presented to and published by the Society. The paper is "The last unpublished notes of J. Willard Gibbs" (*Proc. Amer. Philos. Soc.*, 1961, 105, 545–588). Jerzy Konorski of the Nenski Institute in Warsaw was elected a Foreign Associate. The other Foreign Associates in psychology are Bartlett, Michotte, and Piéron.

No papers in the field of psychology were read. The psychologists in attendance were E. G. Boring, Leonard Carmichael, W. R. Miles, C. P. Richter, R. R. Sears and S. S. Stevens.

Harvard University EDWIN G. BORING

THE 1963 MEETING OF THE NATIONAL ACADEMY OF SCIENCES

The 1963 Meeting of the National Academy of Sciences was held at the building of the Academy in Washington on April 22-24, 1963. Walter R. Miles presented a report on "Chimpanzee behavior: Removal of a foreign body from a companion's eye." C. H. Graham, Yun Hsia, and F. F. Stephan read on "Visual discrimination of subject with acquired unilateral tritanopia." R. W. Gerard, T. J. Chamberlain, and G. H. Rothschild reported on "RNA in learning and memory." William Shockley presented a paper on "Proposed important mental tools for scientific thinking at the high-school level."

William K. Estes of Stanford University was elected a member of the Academy. The membership of the Section, diminished by the death of R. S. Woodworth on July 4, 1962, thus continues at 24.

The psychologists in attendance were E. G. Boring, Leonard Carmichael, Robert Galambos, C. H. Graham, E. R. Hilgard, Heinrich Klüver, D. B. Lindsley, W. R. Miles, G. A. Miller, N. E. Miller, Carl Pfaffmann, C. P. Richter, L. A. Riggs, and S. S. Stevens.

Harvard University EDWIN G. BORING

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Freud's Concept of Repression and Defense: Its Theoretical and Observational Language. By PETER MADISON. Minneapolis, University of Minnesota Press. 1961. Pp. 205. $4.75.

This is an important work which should be studied carefully by all interested in motivational theory and research, or in the clinical application of psychoanalysis. Spurred on by finding that psychoanalytic practitioners seemed to mean more by repression than amnesia, and by lack of clarity as to what else was included in the term, Madison set himself the task of analyzing Freud's writings thoroughly to find out what Freud meant by repression. The first and most extensive part of the book contains a carefully and extensively documented report of his findings. Perhaps other students of Freud's writings could find fault with his conclusions, but to this non-expert it seems difficult to see how the conclusions can be refuted. The quotations from Freud are lengthy enough so that they cannot be said to be out of context, and each point is backed up by numerous quotations drawn from many different stages in the development of Freud's theory. The second part of the book consists of a formal summary of the theory, using a theoretical language based upon Freud's own usage. The final part of the book consists of detailed suggestions on how the theoretical constructs could be translated into an "observational" language which would make it possible to count and measure the manifestations of repression and defense, and thus to validate the theory.

The discussion in this book is so closely reasoned, that it is dangerous to try to condense it further. All that will be attempted here will be to give some idea of the *kind* of ideas it contains. A principal theme of the book is that Freud nearly always used the term *repression* to cover a large group (but not all) of the defenses, and that the concept is not identical with any of these defenses themselves. Instead, Freud had an underlying concept of a process which was reflected in these varied concrete manifestations.

As Madison puts it, "Just as 'heat' refers theoretically to hypothesized atomic vibrations or molecular motion, so 'repression' and 'defense' refer to the hypothesized interplay of force and counterforce (instincts and anticathexis) in psychic life. Just as heat energy becomes manifest in the volumetric expansion of solids or liquids, in the alteration of their electrical resistivity, in the color changes in metals, or in the psychological sensation of hotness to touch, so the interplay of instinct and anticathexis becomes manifest in a wide variety of effects upon psychological processes. Amnesia was simply the historically first described manifestation of this interplay of force and counterforce. Resistances, emotional and behavioral inhibitions, and childhood repressions were eventually seen to be new alternative types of manifestation" (p. 9). Once, in 1926, Freud decided to restrict the term repression to the amnesic phenomena. Madison finds, however, that he did not adhere to this restriction later, and that even in the same book repression was used in the broader theoretical sense.

When it comes to the search for observational indicators of repression and defense, Madison believes that this can be done only within the setting of the clinic. The various attempts to study "repression" in the laboratory have dealt with amnesic phenomena in conditions which involved extremely mild forms of emotional conflict compared to the conflicts with which Freudian theory was concerned and which are found in the clinic. Also they have used *kinds* of motive different from those important in repression theory. Moreover, Madison argures, "While these investigations were often conceived as tests of Freud's theory, his theory, or any theory, was rarely discussed. The investigators themselves seemed very uncertain about exactly what they were trying to measure" (p. 4). One of Madison's reasons for writing this book is to clarify what is to be tested or to be compared with other theoretical formulations. He also feels that post-Freudian psychoanalysis, in spite of many innovations, is still playing variations on a basic theme which is expressed in Freud's own work. Consequently, Madison feels, his task had more than a historical purpose.

The general approach which Madison advocates for developing behavioral indicators of defenses is simple in principle, even though it is complex in application. "The most expert and experienced analysts available would have to agree on the best final definitional statement of each defense and select a series of critical examples that would define different *degrees* of the use of the defense in question. These definitions plus their degree-defining examples would constitute a rating scale for the defense. It would then need to be shown that new independent judges would use the definitional statements and rating scale examples to classify reliably new instances of each defense" (p. 165). Madison is able to give numerous examples of material which he drew from Freud's writings and which would meet the specifications above.

Once these measures of different kinds of defense are developed, they can be used to test the theory. As is evident, Madison disagrees with those who consider Freudian theory to be untestable as a whole. He does agree, however, that there are some defenses which seem to be unmeasurable, and consequently some aspects of the theory which are untestable. For example, he is unable to find observational indicators of the anticathectic forces involved in infantile "primal" repression. He also notes that some of Freud's constructs are detectable only if we assume the correctness of the Freudian rules of symbolism.

Thus, what some might consider the most characteristic and important aspects of Freudian theory are untestable—notably the theory of primal repression, castration trauma, and what Madison calls "symbolic sex." Madison, believes, however, that "actual sex (not symbolic sex) was involved in many of Freud's clinical cases, and especially in those that suggested the theory of repression to him. The theory might be true for these defining cases involving real sex and not be true for other cases without manifest sexual content to which Freud generalized the theory by means of his various translation rules" (p. 191). This represents quite an important change in the concept of repression according to Madison's own account. In summarizing the Freudian position, he states that adult repression depends upon unsuccessful primal repression, and the association of adult sexual or aggressive activities with primal traumas.

It is not clear to this reviewer why the symbolic indicators are ruled out by Madison. It is true that their interpretation depends upon arbitrary rules of translation and that they indicate things which cannot be observed directly. Direct validation of the symbols is thus impossible, but this is also true of other indicators used in science. They can also be validated by the consistency of their behavior with predictions made from theory. The basic difficulty would seem to be not with the use of symbols *per se,* but the looseness of the translation rules and the dependence upon the skill, experience, or bias of the translator. In principle, then, it might be possible to improve the situation and to add symbolic indicators to the mass of other observations which would be used in checking the consistency of the Freudian theory.

In view of the amount of scholarly labor required for Madison's analysis of Freud's theory, it seems ungrateful to suggest additions which would improve the book still further. Perhaps the author can follow with another volume or an extended edition of this one, in which we are told more about (a) the extent to which other psychoanalysts accept or modify various features of Freud's conceptions of repression and defense, and (b) the relations between Madison's analysis of these concepts and the explications given by other commentators and interpreters of Freudian theory. Madison would be justified in claiming exemption from this task in view of the labors already performed, but this volume shows him uniquely qualified to further our education in the manner suggested. T.A.R.

Functions of Varied Experience. Edited by DONALD W. FISKE and SALVATORE R. MADDI. Homewood, Illinois, The Dorsey Press, Inc., 1961, Pp. vi, 501. $7.95.

One of the salient developments of the last ten years in psychology has been the rapid diffusion of interest in a certain research area for which it is difficult to find a suitable name. Terms such as "exploratory behavior," "discrepancy motivation," "curiosity," "collative motivation," "arousal" and "experimental aesthetics" have all been used, but they either take account of some aspects or portions of the area without taking due account of others or they bear the stamp of a particular kind of theoretical interpretation. The area in question clearly belongs to the field of motivation, although it invades many other traditional sub-sections of psychology, such as learning, thinking and physiological psychology. The phenomena that fall within the area are clearly related to one another, but we shall probably have to find out a good deal more about them before we can decide what it is essentially that ties them together and what is the best way of referring to them collectively.

One common thread is the importance of variety of stimulation. The kinds of behavior that belong to this research area are intimately affected by the quantity of variety, novelty, complexity, information, in the environment and many of them are clearly aimed at increasing or decreasing this quantity. Fiske and Maddi have chosen to string a wide diversity of topics, representative of the area under discussion, along this thread of "varied experience." They have marshalled chapters by nine well-known investigators, and have themselves contributed two introductory chapters and two chapters each on their special interests. The book is therefore timely and valuable. Its purpose is to draw attention to a multiplicity of facts that traditional theories of motivation, centered on visceral discomforts, must find it difficult to digest and to raise questions for the stimulation of thinking and research. Their achievement of these purposes is such as to put psychology deeply in their debt.

As one must expect with a collection of such creative and single-minded individuals, the contributions vary widely in subject matter, in quality, and, above all, in degree of relevance to the main theme. On the one hand, we have a scholarly comparative analysis of exploration and play in animals by Welker and a review of alternation behavior by Dember. On the other hand, Kamiya's chapter on drowsiness and sleep and Fiske's on "The inherent variability of behavior" deal with variety of response rather than of stimulation. It is true that the editors have chosen the term "varied *experience*," which they may well feel entitled to apply to both stimulus and response variability. The relations between the two could however, have done with more thoroughgoing discussion. One feels that the editors, when they received the manuscripts from their authors, must have found themselves saddled with more heterogeneity than they had anticipated. In the final chapter, they make a gallant attempt to integrate everything with their conceptual framework as a basis, but the reader is left at the end not with a unified picture of a sharply defined topic but rather with an exhilarating patchwork quilt whose seams look none too solid.

In their conceptual framework, Fiske and Maddi do not avail themselves to the full of the adventurous possibilities offered by this novel, challenging material. They propound eight propositions which are scarcely more than statements summing up empirical findings and are all amply supported by extant data. Why organisms possess the characteristics corresponding to these propositions is the question that must arise at this stage of research, and it can only be answered by relating them to more fundamental theories of behavior or to neurophysiological work or to both, but little is done in these directions.

Other theoretical treatments, such as those put forward by Dember and Earl and by this reviewer, are considered, and their dubious points are rightly indicated. Yet, Fiske and Maddi do not take account of the aims that have inspired these groupings and lie a step or two beyond the aims of the theoretical treatment they prefer. They may also be deemed to have done less than justice to the prospects of modifying and extending existing motivational theories, such as those based on drive and drive-reduction.

For instance, Maddi takes issue with this reviewer's suggestion that exploratory behavior is reinforced through reduction of conflict resulting from the acquisition of information. "Why should the organism approach that which creates conflict? It is especially difficult to understand why this should occur because conflict is usually considered to be a punishing state of affairs," he writes. Yet surely, one of the most fruitful insights of Freud, and one of those most fully corroborated by common experience, is that organisms often have a choice between seeking immediate gratification, which often means incomplete gratification and unfortunate long-term consequences, and delaying gratification, which means suffering, or even increasing, discomfort for a while for the sake of opportunities for fuller and safer gratification that will come later. Freud and many later writers have depicted a predilection for the second alternative as a hallmark of psychological maturity, and the implications of this choice with respect to food-seeking, sexuality and fear have received copious discussion.

One of the major advantages of delaying gratification is the possibility of gathering information that will make later action more efficient, and the Freudians, in particular, have postulated a close link between delay of gratification and "reality testing." If there is any validity in these ideas where other forms of satisfaction are concerned, it is surely plausible that higher organisms willl often find themselves in situations where bewilderments, puzzlements, conflicts and frustrations must temporarily be aggravated if the rewarding experience of overcoming them is to be attained. And surely anybody can think of enough instances from the recreational, aesthetic and intellectual departments of life to convince him that higher organisms often plunge eagerly into these situations.

To sum up then, we have here a book that should be of interest to any psychologist, since no branch of psychology can be unaffected by the issues that it raises. The book is indispensable to any member of the growing band of specialists in this gradually crystallizing research area. It will also make a commendable supplementary text for courses in motivation, since the work it presents is of fundamental importance but too new to be adequately covered by standard textbooks.

University of Toronto D. E. BERLYNE

Growing Old: The Process of Disengagement. By ELAINE CUMMING and WILLIAM E. HENRY. New York, Basic Books Publishing Co., Inc., 1961. Pp. xvi, 293. $6.75.

To many writers, the aged in contemporary society represent a group which is isolated from full participation in community life by society itself. However, in this book Drs. Cumming and Henry claim that instead the aging person is an active participant in the withdrawal rather than a passive pawn subjected to social pressures. This conclusion stems from the study of a stratified sample of physically healthy adults free from major economic worries living in the urbanized area of Kansas City.

On the basis of five extensive interviews with each subject (1½ hours each) conducted at 6-month intervals on approximately 100 men and 100 women aged 50 to 70 yr., the authors develop a "disengagement theory" of aging. The essence of their theory is that aging is characterized by a gradual reduction in the interactions between the aging person and others in the social system he belongs to, and that this withdrawal may be initiated by the individual. In fact, the authors claim that this gradual withdrawal is part of the developmental sequence and that failure to accept this progression is the source of frustration in aging people. This concept opposes the generally accepted idea that rejection of older people by society is the source of feelings of isolation and frustration.

Although the proposed theory deals with the concept of disengagement, the indices derived from ratings, inferences, and estimates obtained from interview data relate to the subject's level of engagement in a social context. The derived indices were: Role Count, Interaction Index, Life Space, Perceived Life Space, Alienation, *F*-Score and Ego Energy. The Role Count was the total number of categories of people (household, relatives, friends, neighbors, fellow workers, specific people, church, organizations) mentioned in the interviews (scores 1 to 9). The Interaction Index was a rating (1 to 5) assigned by two independent reviewers of the interview data based "on the amount of each day spent in situations where others could exercise normative controls over the subject's behavior." The Life Space Measure was a weighted score, based on arbitrary values assigned to specific answers to questions involving the frequency of contacts with other people (household members, relatives, friends, etc.). Other indices were based on projective methods including the Draw-a-Person Test and the *TAT* test.

The data presented undoubtedly show that roles become fewer with age for both men and women, with the loss of work and kinsmen accounting for most of this diminution. Since neither of these factors is ordinarily under the control of the subject, it is difficult to characterize this aspect of disengagement as initiated or even encouraged by the aging person. The Interaction Index which deals with the density of interactions also shows a decrement after the age of 65. One cannot take issue with the statement that "The aging person sees fewer kinds of people, less often and for decreasing periods of time as he grows older," but the conclusion that the aging person actively participates in this restriction will require further investigation before it can be accepted as a description of aging in a significant proportion of the population. Evidence for active participation in the disengagement of this special sample of subjects (no problems of health or income) is often anecdotal in nature, as for example, the statement of one respondent, "I've started divesting myself of things" (p. 73). Other evidence is based on an analysis of changes in goals expressed by the older respondents and the presumed diminution in "Ego Energy," which is inferred from interview material and the responses to the Thematic Apperception Test.

This book is important in three respects. First, it represents the first systematic presentation of at least part of the data obtained in the first large scale research program on the psychological and social aspects of aging in a defined population of community residing subjects. The data obtained from this sample of individuals who were carefully selected to exclude subjects with major problems of health and finances fail to confirm many of the stereotypes of aging people. For example, this group of subjects was not characterized by feelings often attributed to the aged, such as dissatisfaction, frustration, anxiety, or an increase in religious piety. Secondly the authors gave special attention to methods of deriving ratings and scores from interview data and the results of projective techniques. Finally, the authors have developed a psychological theory of aging which stresses developmental sequences. Their hypothesis should stimulate further research.

Baltimore City Hospitals N. W. SHOCK

Variations in Value Orientations. By FLORENCE ROCKWOOD KLUCKHOHN and FRED L. STRODTBECK. Evanston, Illinois, Row, Peterson, 1961. Pp. xiv, 437. $7.75.

This excellent book will interest those concerned with cross-cultural generalizations about values and beliefs. The authors are sociologists who have developed a verbal multiple-choice technique for collecting quantifiable data from informants about corresponding "value orientations" in different cultures and sub-groups. The domains of measurement were selected deductively as involving basic and universal human problems with theoretically orthogonal solutions: time focus (past-present-future), activity focus (doing-being), relation to nature (control-harmony-submission), and the social nexus of decisions (individual-collateral-lineal). Specific problem situations with alternative solutions varying on the above dimensions were presented to appropriate samples of adults in five New Mexico cultures—Mormon, Texan, Navaho, Zuni, and Spanish-American. The book describes relevant cultural data, prior predictions, and detailed statistical analyses.

In addition to gross predictive success, the book displays the surprising power of the technique in differentiating groups on core dimensions. Texans for example were individualistic, future-oriented, doing-oriented, and concerned with controlling nature, whereas their neighbors the Spanish-Americans, were present- and being-oriented, felt subject to nature, and more often than the Texans preferred lineal (parent, elder brother, patrón) sources of decision. Differences of rank-order in subordinate as well as dominant preferences appeared where there were three choices permitted.

An analysis of intercorrelations and factor structure conduced on the few dimensions studied suggests that it would be fruitful to expand the number of dimensions to study correlations both intra- and cross-culturally. It would be unfortunate if the future users of the technique identify it too rigidly with its present content.

It is clear that specific situations chosen for testing present sampling problems, and in this study certain situations showed the impact of culture change more than others, leading to internal variability. Also, in such complex behavior it is rarely possible to interpret alternative solutions solely in terms of a single dimension of contrast. The comments of the informants—unfortunately presented seldom, chiefly for the Navahos—reveal differences in interpretation by informants which could reduce the value of an instrument with only a few items.

No distinction was made between beliefs and values, a distinction long emphasized by psychologists which has proved useful in research not cited by the authors. Further, since efforts are made to relate responses to non-verbal behavior, some of the same difficulties arise as beset prediction from projective and verbal personality tests and attitude questionaries. These difficulties warrant a closer look at response determinants by the authors.

The book raises some important prospects for investigations of concern to social psychologists. One is the analysis of social sub-groupings in terms of dimensions normally only applied to holistic studies of simple cultures. Another is the analysis of interrelations and inconsistencies within societies, and the study of devices used in social organizations—analogous to intra-individual mechanisms—for defense against inconsistencies in values and beliefs. These might include terminological differentiations, role and place insulation, and reduced communication between sub-groups. A third is the relation of value differences to inter-group relations.

SUSAN M. ERVIN

University of California

Industrial Psychology. By B. VON HALLER GILMER with the collaboration of W. J. E. CRISSY, R. M. CYERT, LEONARD W. FERGUSON, ROBERT GLASER, LEE W. GREGG, THOMAS L. HILTON, MYRON L. JOSEPH, HARRY W. KARN, EMANUEL KAY, ROBERT E. KRUG, JAMES G. MARCH, ROBERT B. MILLER, ROBERT M. MORGAN, ROBERT S. RAMSAY, KATHLEEN A. THOMPSON. New York, McGraw-Hill Book Company, Inc. 1961. Pp. xiv, 511. $7.50.

This 511-page book with 21 chapters, including 13 written by 15 collaborators, not only attempts to summarize the research findings of industrial psychology, but to describe the art of the profession and to place it in a social setting. It is an ambitious, forward-looking venture but, unfortunately, it attempts too much in the space available to the authors.

The book is intended, according to the final sentence in Chapter 1, to help the student "acquire some specific understandings, skills and desires which will prepare you to work with others with increasing consideration, understanding and effectiveness." There are excellent chapters on the "Structure of Organization," "Business Operating Procedures," "Labor Management Relations," and the "Industrial Community" which provide a most useful summary of the industrial environment and of the social problems it contains. These, in this reviewer's opinion, provide an orientation that would be useful to all students in an industrial psychology course.

Perhaps it is unfair, in view of the purpose of the text, to critize the book on the grounds that it makes no effort to relate industrial psychology to the general scientific psychological literature and that it makes no assumption that the student has passed at least one introductory course in psychology. These are serious weaknesses. The book may completely discourage the good student who is interested in exploring the industrial implications of the science of psychology. Nor will it appeal to the good student who wants to look at crucial psychological experiments and studies within the industrial environment. The book will fail to stimulate the good student because space limitations do not permit the authors to develop ideas or present problems. Space limitations are imposed partly by the number of pages available (Personnel Selection is summarized in 21 pages) but the effort to make some chapters interesting, practical and common sense is to waste many pages.

The tone of this review is critical because the reviewer believes good undergraduate students should be challenged by the problems of industrial psychology. The senior author had an excellent idea when he brought 15 collaborators together to produce a readable book of broad scope. His text will be very useful to the personnel or training department staff member who wants a book that is encyclopedic in scope. It will be an excellent text book for an extension course where class members are not assumed to have a knowledge of general psychology.

McGill University EDWARD C. WEBSTER

Mental Hygiene: A Text of Readings. Edited by PETER T. HOUNTRAS. Columbus, Ohio, Charles E. Merrill Books, Inc., 1961. Pp. xi, 577.

As the title suggests, this book is intended as a supplementary text for courses in mental hygiene and personal adjustment. Through his experience in teaching such courses, the editor, like all good instructors, has found it beneficial to present his students an opportunity to explore certain concepts more extensively than can be covered in a single basic text. His approach to the study of mental health is eclectic and stresses both theoretical and research contributions from psychology, cultural anthropology, sociology, and related disciplines. The present volume contains a collection of forty such contributions—mostly from original journal articles.

One of Hountras' goals was to "organize in a single place important different points of view and significant findings from the impressive technical literature not likely to be immediately available to any but advanced students of the subject" (p. 1). The question might well be raised as to the advantages of motivating and guiding students to seek out pertinent journal material on their own. On the other hand, in view of the typical undergraduate student's "journalphobia" and aversion to "do-it-yourself" study, the objective of presenting him with a ready-made package of relevant studies may be legitimate.

The organization of the book parallels that of most mental hygiene texts, which enhances its usefulness as a teaching supplement. The list of contributing authors includes such names as Abraham Maslow, Clare Thompson, Frederick Thorne, Arthur Jersild, William Menninger, Carl Rogers, Ruth Benedict, and Percival Symonds. This, in itself, should supply "face" validity for the authoritativeness of the book. Indeed, some of the contributions, *e.g.* Maslow's "A Theory of Human Motivation" might almost be considered classics in the field. It was a little disappointing, however, to note that over a third of the chapters (14 out of 40, and including 4 of the 5 articles on adolescence) are of pre-1950 vintage. Unfortunately, the flavor of some of these does not improve with age. An example would be Finger's "Sex Beliefs and Practices among Male College Students." This study is based on questionnaires obtained from 111 unmarried, advanced psychology students from one school during the years 1943-44. The results derived from this limited, homogenous sampling taken nearly twenty years ago would seem to have questionable relevance to the college population of today.

Despite this limitation, this reviewer is inclined to agree that the majority of these selections meet Hountras' aim to "provoke animated discussion, making enjoyable the task of learning" (p. 2). Any medium which furthers this end is worthy of merit.

Los Angeles State College ROBERT T. LEWIS

Contemporary Issues in Thematic Apperception Methods. Edited by JEROME KAGAN and GERALD S. LESSER. Springfield, Illinois, C. C Thomas, 1961. Pp. xiv 328. $12.00.

Like so many books that attempt to report papers and discussion presented at an informal meeting, the contributions of participants in the Yellow Springs Conference vary greatly in quality. Eight formal papers are presented, each followed by edited remarks and group discussion. In an effort to preserve the spontaneity of the original remarks, the editors have retained more of the discussion than is really necessary, occasionally resulting in a rather tiresome, disjunctive round of superficial comments.

The opening paper by Robert Holt starts with a well-written analysis of the differences between fantasy, daydreams or reveries, and *TAT* stories. These are subtle but highly significant points as far as any interpretation is concerned. Holt goes on to present an extension of Rapaport's work on a psychoanalytic theory of thinking and its application to thematic analysis. His final point that *TAT* stories are really not fantasies is well taken.

The second paper by Richard Lazarus challenges the earlier, direct-expression theory which states that strong motives are directly expressed in *TAT* stories. As an alternative he proposes two principles: (1) a substitutive principle which maintains that there will be a direct relationship between need states and fantasy only when the need has been blocked from motoric discharge; and (2) a defensive principle which states that under excessively high levels of need, various ego-control processes serve to alter or eliminate the cognitive expression of the need, thereby reducing the accompanying anxiety. While John Atkinson then replies in an attempt to defend the current Atkinson-McClelland position, Lazarus' criticisms are highly germane to the fundamental dilemma that runs through the entire conference as an unresolved issue. To what extent and under what conditions do *TAT* stories reflect the motives or needs of the story-teller? No one provides a satisfactory answer.

Other major contributors include Joseph Veroff who generalizes with undue enthusiasm from minor group trends obtained in a nationwide sample survey employing several *TAT* cards; Seymour Feshbach who employs a mediating stimulus model in a discussion of conflict and fantasy behavior; and Edwin Shneidman who develops a novel; though highly speculative, "psycho-logic" for relating thought patterns to other personality characteristics. Jerome Kagan presents some specific studies dealing with stylistic variables in fantasy behavior, and Bernard Murstein reviews systematically the importance of the stimulus as a determinant of *TAT* stories. The final paper by Douglas Kenny extends still further the reappraisal of stimulus factors in the *TAT*.

University of Texas WAYNE H. HOLTZMAN

An Introduction to Social Psychology. By B. KUPPUSWAMY. New York: Asia Publishing House, 1961. Pp. xv, 572. $5.50.

This textbook, published in India and written by an Indian for Indian students, provides a competent and intersting survey of the field of social psychology. The book is divided into five parts: "Introduction," "Social Interaction," "Socialization," "Group Process," and "Applications." There is an initial emphasis on social behavior at the individual level of analysis (*e.g.* development of social behavior, social perception, roles and status, language, motivation, and attitudes) with a later emphasis on group-level analysis of social behavior (*e.g.* the nature and formation of groups, mass behavior, and leadership). The final section on applications is devoted to three major areas: juvenile delinquency, industrial life, and revolution, war and peace. Like many introductory textbooks of social psychology, this one is descriptive, with little critical evaluation of the concepts, principles, and experiments presented.

Professor Kuppuswamy states that he wrote this book in order to make the principles of social psychology more meaningful to Indian students. From this viewpoint, the text should be highly successful because of the extensive use of illustrations from Indian life and Indian problems. It is conceivable that a major contribution of this textbook would be to promote the growth of social psychology in India.

Instructors in this country may wish to use this text for supplementary reading in social psychology courses. The book contains descriptions of studies done in India which are not generally available. Social psychological concepts are examined in the context of current Indian problems and in relation to the traditional philosophies of India. This book would have less appeal as a basic textbook because Professor Kuppuswamy draws heavily from the writings of American social psychologists; there is little new in the way the field of social psychology is conceptualized to appeal to instructors and students in this country. Many references are made to European editions of books that were originally published in this country. (Incidentally, there are some incorrect reference notations and some incorrect references.) It should also be noted that the book is printed on poor quality paper and that it is not likely to stand up under a fair amount of use.

In summary: The content of the text certainly meets current standards for introductory textbooks of social psychology. While the Indian illustrations will have special appeal to Indian students, this text is sure to find an interested international audience as well.

Indiana University SEYMOUR BERGER

Leadership and Organization: A Behavioral Science Approach. By ROBERT TANNENBAUM, IRVING R. WESCHLER, and FRED MASSARIK. New York, McGraw-Hill Book Company, Inc., 1961. Pp. xiv, 456. $7.50.

The reader whose interest is excited by this book's promising title will be disappointed in most of his expectations.

There are four sections. Part One, "Leadership and the Influence Process," is entirely hortatory, containing chapters such as "How to Choose a Leadership Pattern." I found these essays repetitious and often unconvincing, with uncertain evidential antecedents.

Part Two is all about "Sensitivity Training," a commercialized adaptation of the Bethel variety of training in group dynamics. Here the reader will find a good description, with excerpts from transcripts, of group sessions in which participants are encouraged to comment upon one another's performance and motives. Unfortunately, the reader will not find satisfactory evidence demonstrating that the technique has any useful effects (aside from the enthusiastic testimony of some participants).

Part Three, "Studies in Organization," shares some of the defects of Part One, but also contains several interesting research reports. In particular, I enjoyed chapters on "The Multi-Relational Sociometric Survey," a useful extension of sociometric methodology; "Problems in the Use of Indirect Methods of Attitude Assessment," a case history of some unanticipated reactions to a study of labor mediation; and "Assessing Organizational Effectiveness," a study of scientific productivity and morale.

Part Four, "Commentaries," contains essays on the rest of the book by George Bach, Robert Dubin, and Lyndall Urwick. Their critical comments are so much at variance with one another that one suspects the book has served as a projective device, enabling each commentator to state his own position.

Virtually all of the chapters have previously been published in business and management journals. The book is one of the McGraw-Hill Series in Management. I sincerely hope that the business community does not give it serious attention. The specialist in industrial and labor relations, or in personnel psychology, may find several useful chapters, but—with the exception of the commentaries of Part Four—he will not need to consult this book, since the useful chapters have already been published in familiar sources such as *Occupational Psychology* or *Sociometry.*

Cornell University LEO MELTZER

Social Behavior: Its Elementary Forms. By GEORGE C. HOMANS. New York, Harcourt, Brace & World, Inc., 1961. Pp. viii, 404. $5.50.

Several criteria can be applied in assessing the contribution of a work to social psychology: (a) utility in providing mechanisms for understanding the observables of human social behavior; (b) generalizability even within the diversity of human societies; (c) compatibility with established knowledge of individual psychology. For what it provides in speaking particularly to the first two points, this work stands as a significant contribution with the sweep of provocative analysis and the flavor of the research literature.

In his earlier book, *The Human Group,* Homans presented a useful framework within which to view groups of various kinds. This newer effort treats some of the same considerations, but reflects the movement of his concern toward the statement of propositions about elementary social behavior. While this suggests the concept of "informal" relationships, Homans feels that term to be too confining, since "elementary social behavior occurs at all times and never lacks form." He goes on to develop a scheme for understanding social behavior, in the broadest sense, as ". . . an exchange of activity, tangible or intangible, and more or less rewarding or costly, between at least two persons." It is only partly true, however, to say that this is simply an "economic model," since both rewards and costs are treated with considerable sophistication and variegated detail.

In evolving this scheme, Homans brings to bear the basics of operant conditioning, and on his own terms succeeds in pointing out their utility for reckoning with common phenomena of everyday social experience. Though psychologists may feel this application is limited, he introduces concepts such as "distributive justice" and "status congruency" which sustain his analysis on other grounds. As he plainly states: "This book pretends to be a complete psychology no more than it pretends to be a complete sociology."

The reviewer finds several distracting structures carried over by Homans, most particularly the treatment of conformity in an essentially normative mold. Because of their comparability, one also misses the development of a bridge with collateral work of Heider on interpersonal perception, and Festinger with dissonance theory. But these considerations do not detract from an undertaking of considerable proportions that clearly succeeds more often than not in achieving its aims. It deserves the wide attention it will no doubt share with its predecessor, *The Human Group.*

University of Buffalo E. P. Hollander

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

Alexander, Franz, *The Scope of Psychoanalysis* 1921-1961: *Selected Papers.* New York, Basic Books, 1962. Pp. xix, 594. $12.50.

Allinsmith, Wesley and Goethals, George W., *The Role of Schools in Mental Health.* New York, Basic Books, 1962. Pp. xiv, 337. $7.50.

Argyris, Chris, *Interpersonal Competence and Organizational Effectiveness.* Homewood, Ill., Richard D. Irwin, Inc., 1962. Pp. xi, 292. $6.50.

Back, Kurt W., *Slums, Projects, and People: Social Psychological Problems of Relocation in Puerto Rico.* Durham, N.C., Duke Univ. Press, 1962. Pp. xiii, 123. $5.00.

Battis, Emery, *Saints and Sectaries: Anne Hutchinson and the Antinomia Controversy in the Massachusetts Bay Colony.* Chapel Hill, N.C., Univ. North Carolina Press, 1962. Pp. xv, 379. $7.50.

Baughman, E. E., and Welsh, G. S., *Personality: A Behavioral Science.* Englewood Cliffs, N.J., Prentice-Hall, 1962. Pp. 566. $7.25.

Berkowitz, Leonard, *Aggression: A Social Psychological Analysis*. New York, McGraw-Hill, 1962. Pp. xv, 361. $7.95.

Bernard, Harold W., *Human Development in Western Culture*. Boston, Allyn and Bacon, Inc., 1962. Pp. xii, 445. $7.50 text, $9.95 trade.

Berne, Eric, *A Layman's Guide to Psychiatry and Psychoanalysis*. New York, Grove Press, 1947. Pp. xxi, 320. $.75. (paperback)

Bettelheim, Bruno, *Dialogues with Mothers*. New York, Free Press of Glencoe, 1962. Pp. 216. $3.50.

Blood, Robert O., Jr. *Marriage*. New York, Free Press of Glencoe, 1962. Pp. xix, 515. $6.00.

Brill, Norman Q., (ed.). *Psychiatry in Medicine*. Berkeley, Calif., Univ. of California Press, 1962. Pp. xvi, 195. $3.95.

Brim, O. G., Glass, D. C., Lavin, D. E., and Goodman, N., *Personality and Decision Processes: Studies in the Social Psychology of Thinking*. Stanford, Calif., Stanford Univ. Press, 1962. Pp. ix, 336. $8.00.

Broad, C. D., *Lectures in Psychical Research*. (Incorporating the Perrott Lectures given in Cambridge University in 1959 and 1960). New York, Humanities Press, 1962. Pp. xi, 450. $10.00.

Bruner, Jerome S., *On Knowing: Essays for the Left Hand*. Cambridge, Mass. Harvard Univ. Press, 1962. Pp. 165. $3.75.

Buhler, Charlotte, Values in Psychotherapy. New York, Free Press, 1962. Pp. xv, 251. $5.25.

Burgemeister, Bessie B., *Psychological Techniques in Neurological Diagnosis*. New York, Harper & Row, 1962. Pp. viii, 248. $7.50.

Cantril, Hadley, *Human Nature and Political Systems*. New Brunswick, N.J., Rutgers Univ. Press, 1961. Pp. xiii, 112. $3.00.

Caprio, Frank S., *Female Homosexuality: A Modern Study of Lesbianism*. New York, Grove Press, 1962 (copyright 1954). Pp. xvii, 334. $.75. (paperback)

Clark, Burton R., *Educating the Expert Society*. San Francisco, Calif., Chandler Publishing Co., 1962. Pp. xi, 301. $2.75. (paperback)

Coleman, Harry L., and Smallwood, Clarence. *Computer Language: An Autoinstructional Introduction to Fortran*. New York, McGraw-Hill, 1962. Pp. xiv, 196. $3.95 text; $4.95 trade.

Davis, F. J., Foster, H. H., Jr., Jeffrey, C. R., and Davis, E. E., *Society and the Law: New Meanings for an Old Profession*. New York, Free Press, 1962. Pp. vi, 488. $6.95.

Dingwall, Eric J., *Very Peculiar People*. New Hyde Park, N.Y., Univ. Books, Inc., 1962. Pp. 223. $6.00.

Dominian, J., *Psychiatry and the Christian*. New York: Hawthorn Books, 1962. Pp. 138. $3.50.

Dry, Avis M., *The Psychology of Jung: A Clinical Interpretation*. New York, John Wiley & Sons, 1962. Pp. xiv, 329. $6.00.

Dudley, Geoffrey A., *Your Personality and How to Use It*. New York, Emerson Books, Inc., 1962. Pp. 160. $2.95.

Essien-Udom, E. U., *Black Nationalism: A Search for an Identity in America*. Chicago, Univ. of Chicago Press, 1962. Pp. xiii, 367. $6.95.

Evans, Luther H., and Arnstein, George E., (eds.). *Automation and the Challenge to Education*. Washington, D.C., National Education Assoc., 1962. Pp. ix, 190. $4.00.

Fish, F. J., *Schizophrenia*. Baltimore, Md., Williams & Wilkins Co., 1962. (copyright John Wright & Sons, Great Britain). Pp. x, 190. $7.00.

Fleming, C. M., *Adolescence: Its Social Psychology*. New York, Grove Press, 1962 (first published 1948). Pp. vii, 262. $2.45. (paperback)

Fodor, Nandor, *New Approaches to Dream Interpretation*. New York, Citadel, 1962 (copyright 1951). Pp. 368. $1.95. (paperback)

Forster, Francis M., *Synopsis of Neurology*. St. Louis, Mo., C. V. Mosby, 1962. Pp. 225. $6.75.

Foss, Martin, *Logic and Existence*. New York, Philosophical Library, 1962. Pp. 240. $4.75.

Gayen, A. K., *et al., Measurement of Achievement in Mathematics: A Statistical Study on Effectiveness of Board and University Examinations in India.* Calcutta, India, Indian Institute of Technology, Pp. vii, vii, 176, xiv.

Georgopoulos, Basil S., and Mann, Floyd C., *The Community General Hospital.* New York, Macmillan, 1962. Pp. xiv, 693. $12.50.

Greenberger, Martin, (ed.), *Management and the Computer of the Future.* Cambridge, Mass., MIT Press, and New York, John Wiley & Sons, 1962. Pp. xxvi, 340. $6.00.

Guttmacher, Manfred, *The Mind of the Murderer.* New York, Grove Press, 1962. (First published 1960) Pp. xii, 239. $.60. (paperback)

Haire, Mason, (ed.), *Organization Theory in Industrial Practice.* New York, John Wiley & Sons, 1962. Pp. x, 173, $5.75.

Harrison, Jane E., *Epilegomena to the Study of Greek Religion and Themis: A Study of the Social Origins of Greek Religion.* New Hyde Park, N.Y., Univ. Books, Inc., 1962. Pp. lvi, 600. $10.00.

Havighurst, R. J., et al, *Growing Up in River City.* New York, John Wiley & Sons, 1962. Pp. xiii, 189. $4.50.

Hirt, Michael, (ed.), *Rorschach Science: Readings in Theory and Method.* New York, Free Press of Glencoe, 1962. Pp. ix, 438. $8.50.

Homans, George Caspar, *Sentiments and Activities.* New York, Free Press of Glencoe, 1962. Pp. 326. $6.50.

Hyman, H. H., Wright, C. R., and Hopkins, T. K., *Applications of Methods of Evaluation: Four Studies of the Encampment for Citizenship.* Berkeley, Calif., Univ. of California Press, 1962. Pp. xii, 396.

Jennings, Eugene E., *The Executive: Autocrat, Bureaucrat, Democrat.* New York, Harper & Bros., 1962. Pp. xiv, 272. $5.75 trade, $4.50 text.

Jersild, A. T., Lazar, E. A., and Brodkin, A. M., *The Meaning of Psychotherapy in the Teacher's Life and Work.* New York, Bureau of Publications, Teachers College, Columbia Univ., 1962. Pp. 151. $3.95.

Johnston, N., Savitz, L., and Wolfgang, M. E., (eds.), *The Sociology of Punishment and Correction.* New York, John Wiley & Sons, 1962. Pp. xi, 349. $4.25 paper,, $6.50 cloth.

Jones, Richard M., *Ego Synthesis in Dreams.* Cambridge, Mass., Schenkman Publishing Co., 1962. Pp. 100. $2.45 paper, $4.50 cloth.

Knapen, Marie-Therese, *L'Enfant Mukongo.* Louvain, Editions Nawelaerts, 1962. Pp. 204. FB 170.

Kosa, John, *Two Generations of Soviet Man: A Study in the Psychology of Communism.* Chapel Hill, N.C., Univ. of North Carolina Press, 1962. Pp. viii, 214. $5.00.

Kraepelin, Emil, *One Hundred Yearsof Psychiatry.* New York, Citadel Press, 1962. Pp. 163. $1.75. (paperback).

Krech, David, Crutchfield, Richard S., and Ballachey, Egerton L., *Individual in Society: A Textbook of Social Psychology.* New York, McGraw-Hill, 1962. Pp. 564. $7.95.

Krich, A. M., (ed.), *The Homosexuals as Seen by Themselves and Thirty Authorities.* New York, Citadel Press, 1962 (copyright 1954). Pp. 342. $1.95. (paperback).

Laing, R. D., *The Self and Others: Further Studies in Sanity and Madness.* Chicago, Quadrangle Books, Inc., 1962. Pp. xi, 186. $6.00.

Lantz, Herman R., and Snyder, Eloise C., *Marriage: An Examination of the Man-Woman Relationship.* New York, John Wiley & Sons, 1962. Pp. xii, 417. $6.50.

Lennhoff, F. G., *Exceptional Children: Residential Treatment of Emotionally Disturbed Boys at Shotton Hall.* New York, Taplinger Publishing Co., 1962. Pp. 201. $4.95.

Lesche, Carl, *A Metascientific Study of Psychosomatic Theories and Their Application in Medicine.* New York, Humanities Press, 1962. Pp. 64. $1.75. (paperback)

Lichter, S. O., Rapien, E. B., Seibert, F. M., and Skiansky, M. A., *The Drop-Outs.* New York, Free Press, 1962. Pp. xiii, 302. $5.50.

LINDNER, ROBERT, *Prescription for Rebellion.* New York, Grove Press, 1962 (copyright 1952). Pp. 305, v. $1.95. (paperback)

LINDNER, ROBERT, *Rebel Without a Cause: The Story of a Criminal Psychopath.* New York, Grove Press, 1962 (copyright 1944 by Grune & Stratton). Pp. xvi, 335. $.75. (paperback)

MARTIN, J. P., *Offenders as Employees (Cambridge Studies in Criminology).* New York, St. Martin's Press; London, Macmillan & Co., Ltd., 1962. Pp. xiii, 178. $7.00.

MAURER, DAVID W., and VOGEL, VICTOR H., *Narcotics and Narcotic Addiction,* 2nd ed. Springfield, Ill., Charles C Thomas, 1962. Pp. xii, 339. $9.00.

MCCANN, RICHARD V., *The Churches and Mental Health.* New York, Basic Books, 1962. Pp. x, 278. $6.00.

MCGOWAN, J. F., and SCHMIDT, L. D., *Counseling: Readings in Theory and Practice.* New York, Holt, Rinehart & Winston, 1962. Pp. xiv, 623.

MCKENZIE, JOHN G., *Guilt, Its Meaning and Significance.* London, George Allen & Unwin, Ltd., 1962. Pp. 192. 21s.

MERENESS, DOROTHY, and KARNOSH, LOUIS J., *Essentials of Psychiatric Nursing,* 6th ed., St. Louis, Mo., C. V. Mosby Co., 1962. Pp. 312. $5.50.

MEYNARD, BENJAMIN, *The Nature of Ego: A Study.* New York, Philosophical Library, 1962. Pp. 211. $4.50.

MULIAN, HUGH, and ROSENBAUM, MAX, *Group Psychotherapy: Theory and Practice.* New York, Free Press of Glencoe, 1962. Pp. xvi, 360. $5.95.

MUUSS, ROLF E., *Theories of Adolescence.* New York, Random House, 1962. Pp. 184. $1.95. (paperback)

NURNBERGER, JOHN I., *Diagnostic Signs and Symptoms of Emotional Disorders.* Grand Rapids, Mich., Aquinas College, 1961. Pp. 40. $.75. (paperback)

NUTTIN, JOSEPH, *Psychoanalysis and Personality: A Dynamic Theory of Normal Personality.* New York, New American Library, 1962 (copyright 1953). Pp. xvi, 332. $.75. (paperback)

ORTMAN, E. J., *Philosophy of Teaching.* New York, Philosophical Library, 1962. Pp. 398. $4.75.

PACKARD, VANCE, *The Pyramid Climbers.* New York, McGraw-Hill, 1962. Pp. x, 339. $5.00.

PACKARD, VANCE, *The Status Seekers: An Exploration of Class Behavior in America and the Hidden Barriers that Affect You, Your Community, Your Future.* New York, David McKay Co., 1959. Pp. viii, 376. $4.50.

PARKER, BEULAH, *My Language is Me: Psychotherapy with a Disturbed Adolescent.* New York, Basic Books, 1962. Pp. viii, 397. $8.50.

PATKA, FREDERICK, *Existentialist Thinkers and Thought.* New York, Philosophical Library, 1962. Pp. 170. $4.75.

PITTMAN, D. J., and SNYDER, CHARLES R., (eds.), *Society, Culture, and Drinking Patterns.* New York, John Wiley & Sons, 1962. Pp. xvii, 616. $9.75.

POWDERMAKER, HORTENSE, *Copper Town: Changing Africa. The Human Situation on the Rhodesian Copperbelt.* New York, Harper & Row, 1962. Pp. xxiii, 391. $7.95.

PRIMROSE, E. J. R., *Psychological Illness: A Community Study.* Springfield, Ill., Charles C Thomas, 1962. Pp. vii, 60. $3.50.

REIK, THEODOR, *Ritual: Four Psychoanalytic Studies.* New York, Grove Press, 1962 (copyright 1946). Pp. 367. $2.45. (paperback)

REIK, THEODOR, *Jewish Wit.* New York, Gamut Press, 1962. Pp. 246. $2.45. (paperback)

RITEY, HECTOR J., *The Human Kingdom: A Study of the Nature and Destiny of Man in the Light of Today's Knowledge.* New York, University Publishers, 1962. Pp. 498. $10.00.

ROBINSON, MARIE N., *The Power of Sexual Surrender.* New York, New American Library, 1962 (copyright 1959). Pp. viii, 192. $.60. (paperback)

ROSENFELS, PAUL, *Psychoanalysis and Civilization.* New York, Libra Publishers, Inc., 1962. Pp. 222. $3.95.

SAGARIN, EDWARD, *The Anatomy of Dirty Words.* New York, Lyle Stuart, 1962. Pp. 220. $4.95.

SALZMAN, LEON, and MASSERMAN, JULES, (eds.), *Modern Concepts of Psychoanalysis.* New York, Citadel Press, 1962. Pp. 210. $4.75 cloth, $1.95 paper.

SARASON, IRWIN G., (ed.), *Contemporary Research in Personality.* Princeton, N.J., D. VanNostrand Co., 1962. Pp. xii, 411. $7.00.

SARASON, SEYMOUR B., DAVIDSON, KENNETH, and BLATT, BURTON, *The Preparation of Teachers: An Unstudied Problem in Education.* New York, John Wiley & Sons, 1962. Pp. xv, 124. $3.95 cloth, $1.95 paper.

SARNOFF, IRVING, *Personality Dynamics and Development.* New York, John Wiley & Sons, 1962. Pp. xv, 572. $7.75.

SATERSTROM, MARY HORKHEIMER, and STEPH, JOE A., (eds.), *Educators Guide to Free Guidance Materials.* Randolph, Wisc., Educators Progress Service, 1962. Pp. xv, 253. $6.50 (paperback).

SCHMITT, FRANCIS O., (ed.), *Macromolecular Specificity and Biological Memory.* Cambridge, Mass., M.I.T. Press, 1962. Pp. 119. $3.00.

SHERIF, MUZAFER, (ed.), *Intergroup Relations and Leadership: Approaches and Research in Industrial, Ethnic, Cultural, and Political Areas.* New York, John Wiley & Sons, 1962. Pp. xiv, 284. $5.95.

SINCLAIR, UPTON, *Mental Radio.* Springfield, Ill., Charles C Thomas, 1930 (rev. 2nd printing 1962). Pp. xv, 237. $8.50.

SMITH, DONALD C., *Personal and Social Adjustment of Gifted Adolescents.* Washington, D.C., Council for Exceptional Children, Research Monograph, Series A, No. 4, 1962. Pp. 64. $2.00. (paperback)

SMITH, ERNEST A., *American Youth Culture: Group Life in Teenage Society.* New York, Free Press of Glencoe, 1962. Pp. vii, 264. $5.00.

SMITH, M. G., *West Indian Family Structure.* Seattle, Wash., Univ. of Washington Press, 1962. Pp. vii, 311. $6.00.

SROLE, LEO *et al., Mental Health in the Metropolis: The Midtown Manhattan Study, Vol. 1.* New York, McGraw-Hill, 1962. Pp. xii, 428. $9.95.

STEPHENS, WILLIAM N., *The Oedipus Complex: Cross-Cultural Evidence.* New York, Free Press of Glencoe, 1962. Pp. x, 273. $5.00.

STOODLEY, BARTLETT H., (ed.), *Society and Self: A Reader in Social Psychology.* New York, Macmillan Co., 1962. Pp. 713. $7.50.

STRUPP, HANS H., and LUBORSKY, LESTER, (eds.), *Research in Psychotherapy. Vol. II.* Washington, D.C., American Psychological Assoc., 1962. Pp. ix, 342. $3.50.

TANNENBAUM, ABRAHAM J., *Adolescent Attitudes Toward Academic Brilliance.* New York, Teachers College, Columbia Univ., 1962. Pp. x, 100. $3.50.

THORNTON, HARRY and AGATHE (in collaboration with LIND, A. A.), *Time and Style: A Psycho-Linguistic Essay in Classical Literature.* London, Methuen & Co., Ltd., 1962. Pp. xii, 131. 30s.

TORRANCE, E. PAUL, *Guiding Creative Talent.* Englewood Cliffs, N.J., Prentice-Hall, 1962. Pp. xi, 278. $5.00 text, $6.65 trade.

TOMKINS, SILVAN S., *Affect Imagery Consciousness. Vol. I, The Positive Affects.* New York, Springer Pub. Co., Inc., 1962. Pp. xi, 522. $8.00.

TURBAYNE, COLIN M., *The Myth of Metaphor.* New Haven, Conn., Yale Univ. Press, 1962. Pp. viii, 224. $6.00.

UHLER, ALFRED, *Discover Your Real Assets.* New York, Citadel Press, 1962. Pp. 157. $3.95.

VALENTINE, MAX, An Introduction to Psychiatry, 2nd ed. Edinburgh and London, E. & S. Livingstone, Ltd., 1962. Pp. viii, 316. $5.00.

VOILAND, ALICE L., and associates, *Family Casework Diagnosis.* New York, Columbia Univ. Press, 1962. Pp. ix, 369. $8.50.

WILNER, D. M., WALKLEY, R. P., PINKERTON, T. C., and TAYBACK, M., *The Housing Environment and Family Life: A Longitudinal Study of the Effects of Housing on Morbidity and Mental Health.* Baltimore, Md., Johns Hopkins Press, 1962. Pp. xxii, 338. $7.50.

YATES, AUBREY J., *Frustration and Conflict.* New York, John Wiley & Sons, 1962. Pp. x, 236. $5.00.

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

VOL. LXXVI SEPTEMBER, 1963 No. 3

RESISTANCE TO EXTINCTION IN THE PIGEON AFTER PARTIALLY REINFORCED INSTRUMENTAL TRAINING UNDER DISCRETE-TRIALS CONDITIONS

By WILLIAM A. ROBERTS, DONALD H. BULLOCK, and
M. E. BITTERMAN, Bryn Mawr College

In two recent experiments, the behavior of the pigeon was studied under conditions in which fish and rat seem to function differently. The results obtained were similar to those for the rat in one case and to those for the fish in the other. An experiment on reversal-learning showed progressive improvement in the pigeon, a familiar phenomenon in the rat, but difficult to demonstrate in the fish.[1] An experiment on probability-learning showed matching in the pigeon, now a well-established phenomenon in the fish, but difficult to demonstrate in the rat.[2] These results for the pigeon have a number of interesting implications, of which it may be sufficient, for purposes of illustration, to specify one: The discrepant results for rat and fish cannot be understood in terms of a single dimension of difference. That is, the set of factors which makes for maximizing in the rat where the fish matches is not the same as the set of factors which is responsible for progressive improvement in habit reversal; if they were the same, the pigeon would behave either like the rat or like the fish in both types of experiment.

A third type of experiment in which rat and fish seem to function

---

* Received for publication May 15, 1962. This work was supported by the Office of Naval Research under terms of Contract Nonr 2829(01) with Bryn Mawr College. It was done during Dr. Bullock's tenure of a special USPHS research fellowship.

[1] D. H. Bullock and M. E. Bitterman, Habit reversal in the pigeon, *J. comp. physiol. Psychol.*, 55, 1962, 958-962.

[2] Bullock and Bitterman, Probability matching in the pigeon, this JOURNAL, 75, 1962, 634-639.

differently has to do with the effect of partial reinforcement on resistance to extinction. It is with this problem that the present work on the pigeon is concerned.

EXPERIMENT I

In the rat, resistance to extinction is increased by partial reinforcement both under free-operant and discrete-trials conditions. When the first experiment to be reported here was begun, this so-called paradoxical effect had been demonstrated in the pigeon under free-operant conditions,[3] but not under discrete-trials conditions, and work with the fish which then was in progress suggested that the two procedures did not necessarily give the same result; after a series of experiments with discrete trials which had shown greater resistance to extinction in the fish after *consistent* reinforcement,[4] a free-operant experiment was performed which did give the paradoxical effect.[5] A discrete-trials experiment with the pigeon clearly was required.

It is known now, of course, that the different results found for the fish under the two conditions were due, not to some fundamental difference between the two conditions, but to a confounded difference in frequency of reinforcement; when number of reinforcements rather than number of training trials is matched, the fish shows greater resistance to extinction after partial reinforcement even under discrete-trials conditions.[6] Since the effect appears in the rat with matched trials as well as with matched reinforcements, the discrepant results for the fish which are obtained in matched-trials experiments may be due to a difference in the way in which rat and fish are affected, not (as once was thought) by inconsistency of reinforcement, but by frequency of reinforcement. From this point of view, the discrepant results for partial reinforcement may reflect the same underlying functional difference as the discrepant results obtained in experiments in probability-learning, and in one other type of experiment (not yet mentioned) which also seems to reflect a differential sensitivity to frequency of reinforcement: Over a series of extinctions separated by retraining sessions, the rat's resistance to extinction falls progressively to a low level which is independent of the number of interpolated retraining trials; under like conditions, the fish's resistance also falls progressively, but independence of the amount of interpolated reinforcement does not seem to be achieved.[7] It will be interesting then to examine the behavior of the probability-matching pigeon in an experiment of this sort as well as in a matched-trials experiment in partial reinforcement.

---

[3] W. O. Jenkins, H. McFann, and F. L. Clayton, A methodological study of extinction following aperiodic and continuous reinforcement, *J. comp. physiol. Psychol.*, 43, 1950, 155-167.

[4] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this JOURNAL, 72, 1959, 184-199; Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. comp. physiol. Psychol.*, 53, 1960, 169-172; Wodinsky and Bitterman, Resistance to extinction in the fish after extensive training with partial reinforcement, this JOURNAL, 73, 1960, 429-434.

[5] R. C. Gonzalez, R. M. Eskin, and M. E. Bitterman, Extinction in the fish after partial and consistent reinforcement with number of reinforcements equated, *J. comp. physiol. Psychol.*, 55, 1962, 381-386.

[6] Gonzalez, Eskin, and Bitterman, *op. cit.*, 381-386; Further experiments on partial reinforcement in the fish, this JOURNAL, 76, 1963, 366-375.

Another difference between rat and fish which appears in experiments on partial reinforcement is related to pattern rather than to frequency of reinforcement. The behavior of rats trained with alternating reinforcement and nonreinforcement is distinguishable clearly from that of rats which also are reinforced 50% of the training trials but with reinforced and unreinforced trials scheduled according to Gellermann's rules.[8] In acquisition, the alternating rats develop a tendency to respond more rapidly after nonreinforcement than after reinforcement, which the Gellermann rats do not, and, in extinction, the Gellermann rats show greater resistance.[9] In the fish, by contrast, the two patterns of reinforcement do not have distinguishable effects. The Gellermann animals, like the alternating animals, come to respond more rapidly after nonreinforcement than after reinforcement, and there is no difference in resistance to extinction.[10] The inclusion of an alternating group in the present experiment makes it possible to study the effects of patterning in the pigeon.

Subjects. The *S*s were 21 experimentally naïve White Carneau pigeons, all males, obtained from the Palmetto Pigeon Farm. They were housed individually in a temperature-controlled colony room under conditions which provided constant access to water and grit but only limited access to grain. Throughout the experiments, the animals were maintained at 75% of their satiated body-weights.[11]

Apparatus. For preliminary training, a specially built pecking panel was used. It consisted of a row of tissue-paper covered holes in a metal plate through which the animals could peck for food. The experimental unit was a metal panel, containing a single pigeon-key and a magazine, which was set into a picnic chest. The only sources of illumination in the chamber were the key-light, which came on at the start of each trial and was turned off by response, and the magazine-light, which was on during reinforcement; the chamber was dark during the intertrial interval. The sequence of trials was programmed automatically. The task of *E* was (1) to determine, by throwing a toggle-switch, whether the response on a given trial would be reinforced; (2) to read response-latencies from a Standard Electric timer; and (3) to reset the timer after each reading.

---

[7] Gonzalez, Eskin, and Bitterman, Alternating and random partial reinforcement in the fish, with some observations on asymptotic resistance to extinction, this JOURNAL, 74, 1961, 561-568.

[8] L. W. Gellermann, Chance orders of alternating stimuli in visual discrimination experiments, *J. genet. Psychol.*, 42, 1933, 206-208.

[9] D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, this JOURNAL, 66, 1953, 57-65; E. J. Capaldi, The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses, *J. comp. physiol. Psychol.*, 51, 1958, 367-371.

[10] Gonzalez, Eskin, and Bitterman, *op. cit.*, this JOURNAL, 74, 1961, 561-568.

[11] D. H. Bullock, W. A. Roberts, and M. E. Bitterman, Techniques of housing and maintaining a large pigeon colony, *J. exp. Anal. Behav.*, 4, 1961, 285-286.

Preliminary training. All *S*s were magazine-trained in the experimental unit. On the first day, they were permitted to eat freely from the lighted magazine with the hopper propped up. On the second day, they were given 1 hr. of magazine-training on a VI-20 sec. schedule, each presentation of the magazine lasting 4 sec. On the third day, there was another 1-hr. session of magazine-training, this one with a VI-60 sec. schedule. Each *S* then was given four or five days of experience with the pecking panel. On the first day, openings were punched through the tissue paper covering the holes in the panel, which made it possible for the animals to eat the grain. On the second day, the openings were smaller, and the animals were required to enlarge them to get at the grain. On subsequent days, there were no openings at all. When placed eventually in the experimental unit, all *S*s showed satisfactory positive transfer, pecking at the key and eating from the magazine.

Experimental procedure. The animals were trained in daily sessions of 20 trials with an intertrial interval of 6 sec. Each trial began with the illumination of the key, at which time the latency-clock was started. The response of the animal turned off the key-light, stopped the clock, and, on reinforced trials, caused the magazine to become available for a period of 4 sec.; on unreinforced trials, the intertrial interval began immediately.

In the first seven days of the experiment proper, all responses were reinforced. On the basis of their performance during this period, the animals were divided into three matched groups of 7 *S*s each—a Consistent Group, a Random Group, and an Alternating Group. The training of the Consistent Group continued as before, but partial (50%) reinforcement was introduced for the other two—for the Random Group, the sequence of reinforced and unreinforced trials followed Gellermann's rules while for the Alternating Group odd trials were reinforced and even trials unreinforced. Extinction began on Day 27. Again, there were 20 trials per day, but all were unreinforced. If on any trial, *S* failed to respond in 30 sec., the trial was terminated; that is, the key-light was turned off and the intertrial interval begun. The criterion of extinction was five successive failures to respond in 30 sec. on any given day.

After its first extinction, each animal was given three days of retraining under the conditions which prevailed for it before extinction (the Consistent animals being reinforced consistently, the Random animals in Gellermann orders, and the Alternating animals alternately). Then each animal was extinguished once more to the same criterion as before, retrained for three days, extinguished again, and so forth, until it had been extinguished ten times.

In the final stage of the experiment, two further sets of observations were made. First, out of interest in the effects of alternating reinforcement and nonreinforcement, the Alternating Group was given 10 more days of training under this condition. Secondly, a study was made of resistance to extinction as a function of the number of days of interpolated retraining. For this purpose, a group of 10 *S*s was constituted of 2 Alternating *S*s, 3 Random *S*s and 5 Consistent *S*s. Each animal, serving as its own control, was extinguished after 0, 3, or 6 days of retraining, with order balanced over the group as a whole. All retraining in this final stage of the experiment was with consistent reinforcement.

Results. It had been planned originally to carry out each extinction in a single session, as in previous work with the fish, but the first pigeon to be extinguished (a member of the Alternating Group) required so many trials that the procedure was abandoned in favor of the series of 20-trial sessions. The pigeon extinguished in a single session was discarded from the experiment. With an animal of the Random Group having died, the data to be considered are those of six Alternating, six Random, and seven Consistent animals.

Learning curves were plotted for each group in terms of mean log latency per trial (the raw time-scores, recorded in units of 0.01 sec., being multiplied by 100 and transformed into logs). The curve for the Consistent Group fell in negatively accelerated fashion to an asymptote of about 1.93. The curves for the partially reinforced groups were, of course, quite similar to that for the Consistent Group during the first seven days, during which all groups were consistently reinforced, then rose abruptly with the introduction of partial reinforcement on the eighth day, and thereafter remained at a higher level than did the curve for the Consistent Group. Both partial curves approached the same asymptotic value, which was about 2.01. Analysis of variance (Lindquist's type I) yields a significant Groups-effect ($F = 3.69$ with 2 and 16 *df*), a significant Days-effect ($F = 4.81$ with 18 and 288 *df*), and an insignificant Days × Groups interaction ($F < 1.00$). Orthogonal comparisons show a significant difference between the Consistent Group and the combined partial groups ($F = 7.27$ with 1 and 16 *df*), and an insignificant difference between the two partial groups ($F < 1.00$).

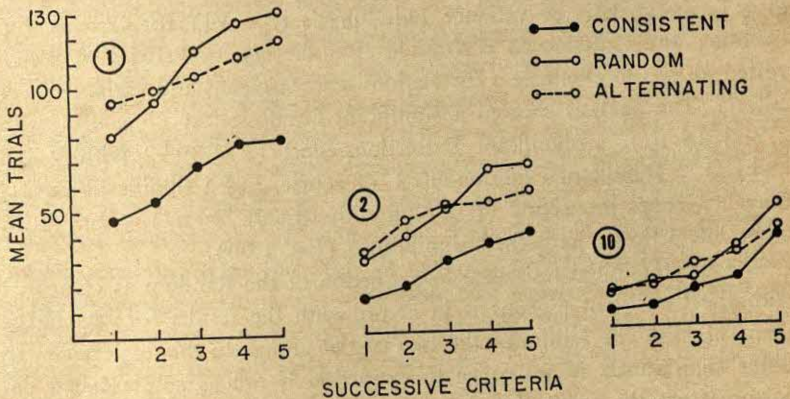


FIG. 1. PERFORMANCE OF THE THREE GROUPS IN THE FIRST, SECOND, AND TENTH EXTINCTIONS OF EXPERIMENT I

In the first part of Fig. 1, performance in the first extinction is plotted in terms of trials to criteria of increasing severity (one, two, three, four, or five successive failures to respond in 30 sec.). This method of plotting was chosen to avoid the arbitrariness associated with the selection of any single criterion. At each criterion, the partial groups showed more resistance to extinction than did the consistent Group, while the difference between the two partial groups was small. Plotted also in Fig. 1 are the results for the second and tenth extinctions, to represent the changes in performance observed over the entire series of extinctions. Each of the three groups shows a negatively accelerated decline in resistance, and the difference between the consistent group and the partial groups tends to disappear. A three-factor analysis of variance (Lindquist's type VI) for Groups, Extinctions, and Criteria was performed after the raw trial-scores were converted to logs to eliminate a correlation between means and variances over criteria. The analysis showed a significant Groups-effect ($F = 4.11$ with 2 and 16 *df*), a significant Extinctions-effect ($F = 39.92$ with 9 and 784 *df*), a significant Criterion-effect, of course, and a significant Extinctions × Groups interaction ($F = 3.54$ with 18 and 784 *df*). None of the other interactions approached significance ($F < 1.00$).

The greater initial resistance to extinction of the Random as compared with the Consistent pigeons is in accord with the results for rats rather than for fish. The failure of the two partial groups in this experiment to differ significantly in resistance to extinction is not as substantial a departure from the results for the rat as it might first seem, because there is no indication that the alternating pattern of reinforcement was discriminated; the latencies of the Alternating animals, like those of the Random animals, were much the same on trials following reinforcement as on trials following nonreinforcement. Capaldi has shown in an experiment with the rat that a differential decline in the resistance to extinction of Alternating animals does not appear until discrimination of the pattern is evidenced by more rapid response on trials following nonreinforcement.[12] Only late in the present experiment did some such evidence of discrimination appear, but by that time the resistance of both partial groups had declined to a low level. The additional training given the Alternating animals after the tenth extinction clearly established the capacity of the animals to discriminate the alternating pattern, and it remains for a further experiment to determine the effect of such a discrimination, once developed, on resistance to extinction in the pigeon. It is conceivable, of course, that the Alternating fish of Gonzalez, Eskin, and Bitterman also failed for many trials to discriminate the pattern as such, if indeed they ever did discriminate it.[13] The tendency which they showed to respond more rapidly after nonreinforced than after reinforced trials, *a tendency which was shown by the Random animals as well,* may have been nothing more than a frustration-effect.

---

[12] Capaldi, *op. cit.,* 367-371.

The tendency which was found in this experiment for the resistance of partially reinforced animals to decline more rapidly over a series of extinctions than that of consistently reinforced animals also requires some comment. A like effect has appeared in a matched-reinforcements experiment with the fish.[14] Quite the opposite results have been obtained, however, in matched-trials experiments with fish[15] and rat,[16] resistance to extinction declining less rapidly in partially reinforced than in consistently reinforced animals. Some further study of this problem is warranted.

In the final stage of the present experiment, when all of the animals had reached what seemed to be a stable level of resistance to extinction with three days of interpolated retraining, amount of interpolated retraining was varied systematically. Resistance to extinction was found to be unaffected. Plotting trials-to-criterion as a function of amount of interpolated retraining gives an essentially flat curve for each of the five criteria of extinction, and a Treatments × *S*s analysis of variance shows the effect of different amounts of retraining to be negligible ($F < 1.00$). In this respect, too, the results for the pigeon resemble those for the rat rather than those for the fish.

EXPERIMENT II

In a second experiment with the pigeon, the effect of partial reinforcement on resistance to extinction was studied with widely spaced trials. In the rat, resistance to extinction is increased by partial reinforcement in spaced as well as in massed training, but the effect is not necessarily independent of distribution in other species. A number of factors conceivably might operate to produce the effect in massed practice which would not operate in spaced practice. Consider, for example, the Hull-Sheffield suggestion that discrimination of the after-effects of reinforcement and nonreinforcement may be responsible for greater resistance in partially reinforced animals under conditions of massed practice.[17] Experiments with spaced trials may show the *insufficiency* of this interpretation at the level of the *rat,* but they do not rule out the possibility that such a discrimination may contribute to the paradoxical effect in the rat, nor that there are some animals in which the paradoxical effect, where it occurs, is to be accounted for entirely in these terms.

---

[13] Gonzalez, Eskin, and Bitterman, *op. cit.*, this JOURNAL, 74, 1961, 561-568.

[14] Gonzalez, Eskin, and Bitterman, *op. cit.*, *J. comp. physiol. Psychol.*, 55, 1962, 381-386.

[15] Wodinsky and Bitterman, *op. cit.*, this JOURNAL, 72, 1959, 184-199.

[16] C. C. Perkins, and A. J. Cacioppo, The effect of intermittent reinforcement on the change in extinction rate following successive reconditionings, *J. exp. Psychol.*, 40, 1950, 794-801.

There are, in fact, some submammalian data which suggest at least that the paradoxical effect appears more readily in massed than in spaced training. Two experiments have been reported with immature turtles of closely related species. In one, with six trials per day and a 30-sec. intertrial interval, partial reinforcement produced greater resistance to extinction.[18] In another, with a 24-hr. intertrial interval, partially and consistently reinforced groups did not differ in resistance.[19] Subsequent work with mature turtles of another species trained with a 24-hr. interval did show somewhat greater resistance after partial reinforcement, but the effect was a small one which appeared only very late in extinction.[20] A preliminary spaced-trials experiment with the goldfish, performed in this laboratory about one year ago, also failed to show the paradoxical effect, even with matched reinforcements. In view of these results, it cannot be taken for granted that the results for the pigeon will be the same in spaced as in massed training.

To maximize the likelihood of finding the paradoxical effect, should such a tendency develop in the pigeon under conditions of spaced practice, two steps were taken in the present experiment. First, a matched-reinforcements design was used. Although the previous experiment gave no indication that the matching of reinforcements is essential for the pigeon, it seemed the safest procedure. Secondly, in addition to a partial group for which reinforced and unreinforced trials were scheduled in Gellermann orders, a second partial group (a so-called Extended Group) was trained with a schedule which involved much longer runs of unreinforced trials. In the fish, an extended schedule of this sort produces much greater resistance than does the Gellermann schedule.[21] Although there is no direct evidence at the level of the rat on the effect of the extended schedule, resistance to extinction increases as probability of reinforcement declines from 50% to lower levels, and decline in probability of reinforcement must bring with it longer runs of unreinforced trials. It is of interest, in any case, to determine whether the extended schedule will affect the pigeon in the same way as it does the fish.

---

[17] V. F. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.,* 39, 1949, 511-526.

[18] N. R. Murillo, J. K. Diercks, and E. J. Capaldi, Performance of the turtle, pseudemys scripta troostii, in a partial-reinforcement situation, *J. comp. physiol. Psychol.,* 54, 1961, 204-206.

[19] Eskin and Bitterman, Partial reinforcement in the turtle, *Quart. J. exp. Psychol.,* 2, 1961, 112-116.

[20] Gonzalez and Bitterman, A further study of partial reinforcement in the turtle, *Quart. J. exp. Psychol.,* 14, 1962, 109-112.

[21] Gonzalez, Eskin, and Bitterman, *op. cit.,* this JOURNAL, 366-375.

Subjects. The Ss were 27 male White King pigeons, all experimentally naïve, obtained from a local pigeon farm. They were maintained at 75% of their satiated body-weights under conditions identical to those used for the Ss of Experiment I.

Apparatus. The apparatus consisted of a runway, with a starting box and goal-box attached at each end, all made of wood. The runway, which was painted gray, was 6 in. wide, 63 in. long, and 10.5 in. high (inside dimensions). The starting box, which was painted black, was 6 in. wide, 17 in. long, and 10.5 in. high. The dimensions of the goal-box, which was painted white, were the same as those of the starting box. The tops of the three units were hinged wooden doors. In the back wall of the goal-box was a round hole (2 in. in diameter) through which S could put its head. Outside the goal-box, just below the hole, was a food-cup, whose contents could not be seen by S until it had put its head through the hole. Mounted above the hole was a white Christmas-tree bulb which illuminated the food-cup and the interior of the goal-box through a rectangular Plexiglas window in the wall of the box. A gooseneck lamp containing a 60-w. bulb, also was placed near the goal-box to provide further illumination of the food-cup and of the goal-box interior. In both end-boxes, there were horizontally sliding doors which were manipulated by E. One separated the starting box and the runway, while the other prevented retracing from the goal-box into the runway.

Two Standard Electric timers were used to provide separate measures of latency and running time. Mounted on the starting box was a microswitch which was thrown when E opened the door of the starting box, thus activating the first clock. In one wall of the runway, 6 in. beyond the door of the starting box, a photocell was mounted which stopped the first clock and activated the second when a beam of light falling upon it was broken by S. The second timer was stopped when S interrupted a light beam falling upon a second photocell, which was mounted 10 in. beyond the entrance of the goal-box. The first clock provided a measure of latency and the second a measure of running time.

The experiment was conducted in a dimly illuminated room, the only sources of light being those which illuminated the goal-area and those which fell on the photocells. An air-conditioner kept the room at a fairly constant temperature and provided a masking noise throughout the experiment.

Procedure. In the first phase of the experiment, each S was given one trial per day for 10 days. There was no preliminary training, the first day of experimental training being the first exposure to the apparatus for each S. Before each trial, S was carried into the experimental room from the adjoining colony room in a special carrying box. E placed S in the starting box, waited for 30 sec., and then opened the starting door. If S did not leave the starting box within 300 sec., E manually guided it into the goal-box. On each trial, S found 5 gm. of grain in the food-cup. On the first trial, practically all Ss required guidance. After entering the goal-box, all the animals found the grain and ate it readily, and, by the end of the first 10 days of training, all Ss were running to the goal-box freely without guidance.

On the basis of their performance in the first phase of the experiment, the animals were divided into three matched groups of nine each. The three groups were treated differently in the second phase of the experiment, which lasted for 50 days. The Gellermann Group was given 50 trials, one each day, half reinforced and half unreinforced in Gellermann orders (the longest run of unreinforced trials being three). On unreinforced trials, the food-cup was empty, and *S* was confined in the goal-box for 60 sec. The Extended Group was treated like the Gellermann Group, except that the sequence of reinforced and unreinforced trials was different, the number of consecutive unreinforced trials being gradually increased to a maximum of eight. (The exact sequence may be seen in Fig. 2.) The Consistent Group, which had 25 reinforced trials, but no unreinforced trials, was divided into two subgroups. The Consistent (G) subgroup (N = 4) had trials on the days when the Gellermann Group was reinforced, while the Consistent (E) subgroup (N = 5) had trials on the days when the Extended Group was reinforced.

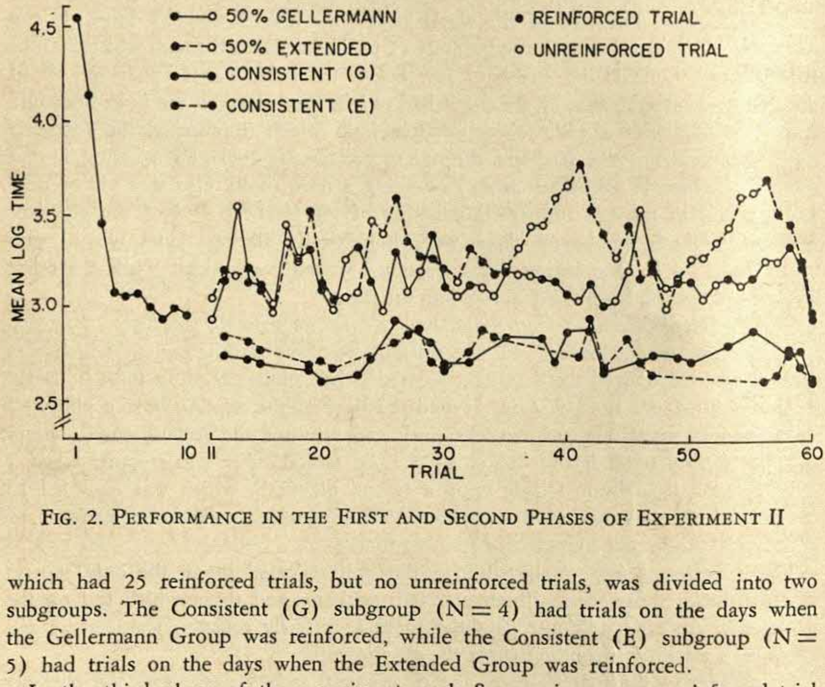


FIG. 2. PERFORMANCE IN THE FIRST AND SECOND PHASES OF EXPERIMENT II

In the third phase of the experiment, each *S* was given one unreinforced trial per day. The procedure on these trials was exactly the same as on unreinforced training trials—the food-cup was empty and *S* was confined in the goal-box for 60 sec. If, on any trial, the animal failed to enter the goal-box in 300 sec., it was removed from the apparatus and the trial terminated. The criterion of extinction was five successive incomplete runs of this kind.

Results. Since the curves for latency and for running time showed essentially the same pattern in all three stages of the experiment, the two time-scores were summed to give a total time-score. (These total scores, originally in uits of 0.01 sec., were multiplied by 100 and transformed into logs.) In the first portion of Fig. 2, the mean performance of all 27 *S*s in the first phase of training is shown. Total time decreases in negatively accelerated fashion until, by Day 4, the curve is close to asymptote. In the second portion of Fig. 2, performance in each of the training conditions of the second phase of the experiment is shown. The Extended curve reflects a persistent tendency of the Extended animals to run faster after reinforced trials (filled circles) than after nonreinforced trials (unfilled circles). The curve for Gellermann animals reflects a like tendency up to about trial 30, after which the pattern disappears. Throughout the second phase, the Consistent animals run more rapidly than do partially reinforced animals. An analysis of variance based on individual mean log times (computed for all trials given each animal in the second phase) yielded a significant Groups effect ($F = 4.35$ with 2 and 24 *df*). The two Consistent subgroups were combined for the purposes of this analysis because they did not differ significantly ($F < 1.00$). Orthogonal comparisons showed a significant difference between the Consistent Group and the combined partial groups ($F = 8.13$ with 1 and 24 *df*), and an insignificant difference between the Gellermann and Extended groups ($F < 1.00$).

Extinction-curves plotted for the first 25 days in the third phase of the experiment are shown in Fig. 3, with the two Consistent subgroups, which again were highly comparable ($F < 1.00$), combined. Resistance to extinction was greater in the partial groups than in the Consistent Group. By Day 25, all *S*s in the Consistent Group had reached the criterion of extinction, while three Gellermann *S*s and four Extended *S*s continued to respond. An analysis of variance (Lindquist's type I) based on log times over the 25 days of extinction shows no significant Groups-effect ($F = 1.11$ with 2 and 24 *df*), but there is a significant Trials-effect ($F = 45.01$ with 24 and 576 *df*) and a significant Trials × Groups interaction ($F = 2.94$ with 48 and 576 *df*), indicating that the Consistent Group and the partial groups extinguished at different rates.

Two features of these results require brief comment. The first is that partial reinforcement increases the pigeon's resistance to extinction in spaced as well as in massed training. There still is no clear evidence, then, of any interaction between consistency of reinforcement and distribution of practice, although, as suggested before, there is some reason to look for such interaction in lower forms. The second fact is that the extended schedule of partial reinforcement does not produce greater resistance to extinction in the pigeon than does the Gellermann. The effectiveness of the extended schedule for the fish may reflect, not a species difference, but an interaction of schedule and distribution, since training was massed in work with the fish. Some experiments on this problem at the level of the rat would be interesting.

The relation of the present findings to those of a previous experiment on partial reinforcement in the pigeon, not heretofore cited, also requires some comment. In that work, with a classical conditioning procedure, no clear evidence of the paradoxical effect could be found.[22] One might conclude from this that certain factors which make for the effect in instrumental training do not operate in classical conditioning, but the findings of a recent experiment on the fish weigh against this interpretation: With matched reinforcements, the fish shows the paradoxical effect in classical as well as in instrumental training.[23] A more reasonable interpretation is that the classical conditioning procedure developed for the pigeon is inadequate. The intense shock used to activate the animal may produce such high resistance to extinction that the effect of partial reinforcement is masked.

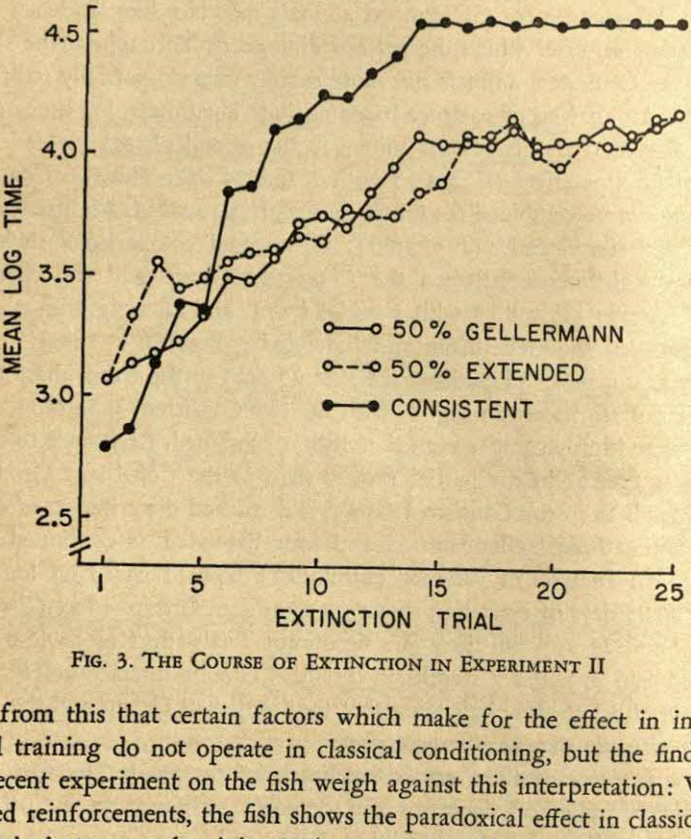


Fig. 3. The Course of Extinction in Experiment II

---

[22] Nicholas Longo, Sandra Milstein, and M. E. Bitterman, Classical conditioning in the pigeon: Exploratory studies of partial reinforcement, *J. comp. physiol. Psychol.*, 55, 1962, 983-986.

[23] Gonzalez, Eskin, and Bitterman, Further experiments on partial reinforcement in the fish, *op. cit.*, this Journal, 366-375.

In general, the behavior of the pigeon in these experiments was quite similar to that which might have been expected of the rat. The only clear discrepancy which appears between the results for the pigeon and those which have been obtained in analogous work with the rat has to do with the differential decline in the resistance to extinction of partially and consistently reinforced animals over a series of extinctions separated by retraining sessions, but there are many differences other than in species between the two studies which could yield the discrepant results, and it is difficult now to assign any weight to the difference. It is only the work on probability-learning, then, which points to any striking functional difference between rat and pigeon.

SUMMARY

Two discrete-trials experiments on partial reinforcement in the pigeon are reported. In the first experiment, a key-pecking apparatus was employed, and three groups of pigeons were trained in massed trials—a Consistent Group (reinforced on every trial), a Random Group (reinforced on 50% of trials in Gellermann orders), and an Alternating Group (reinforced only on odd-numbered trials). After the first extinction, the groups were retrained in the same way as before, extinguished again, retrained again, and so forth, until each had been extinguished ten times. The resistance of the two partial groups was greater in the first extinction than that of the Consistent Group. In subsequent extinctions, resistance declined progressively, more rapidly in the partial groups than in the Consistent Group, and the difference between them tended to disappear. The two partial groups showed comparable resistance at all points in the experiment, but there was no evidence that the alternating pattern had been discriminated until rather late in training, when the resistance of all groups had fallen to a low level. In the final portion of the experiment, resistance to extinction was found to be unaffected by amount of interpolated retraining.

In the second experiment, a runway was used, and trials were widely spaced. Three groups of animals were trained, each given the same number of reinforcements—a Consistent Group (reinforced on all trials), a Gellermann Group (reinforced on 50% of trials), and an Extended Group (reinforced on 50% of trials but with a schedule which involved much longer runs of nonreinforcements than did the Gellermann). Resistance to extinction was greater in the partial groups than in the Consistent, but there was no difference between the two partial groups.

These results for the pigeon are compared with those obtained in analogous experiments with fish and rat.

FURTHER EXPERIMENTS ON PARTIAL REINFORCEMENT IN THE FISH

By R. C. GONZALEZ, ROCHELLE M. ESKIN, and M. E. BITTERMAN,
Bryn Mawr College

The first experiments on partial reinforcement in the fish gave results quite different from those obtained in analogous experiments with the rat.[1] The fish was the African mouthbreeder (*Tilapia macrocephala*), the situation was an instrumental one (striking at a target for a pellet of food), and partial reinforcement was defined as nonreinforcement on half the trials, with Gellermann's rules governing the sequence of reinforced and unreinforced trials. Initial resistance to extinction was found to be *less* after partial than after consistent reinforcement, whether practice was massed or distributed, whether few or many training trials were given, and whether the prevailing level of hunger was high or low. The paradoxical increase in resistance to extinction after partial reinforcement which is so well known in the rat appeared only when partially and consistently reinforced fish were subjected to a series of extinctions separated by retraining sessions; resistance declined progressively in both groups, but that of the consistent group declined more sharply and reached a lower terminal level.[2]

Number of trials or number of (primary) reinforcements—but not both—may be equated in comparing partial and consistent reinforcement. Partially reinforced rats show greater resistance to extinction whether trials or reinforcements are equated, although it is the equated-trials effect which presents the more difficult interpretative problem; the equated-reinforcements effect may be accounted for in terms of secondary reinforcement on the additional trials given the partial animals. In the first experiments on the fish, trials were equated, with the results already described. In subsequent experiments, however, reinforcements were equated, with

---

* Received for publication November 1, 1961. This work was supported by Grant No. M-2857 from the National Institute of Mental Health.

[1] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this JOURNAL, 72, 1959, 184-199; Resistance to extinction in the fish after extensive training with partial reinforcement, this JOURNAL, 73, 1960, 429-434; Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. comp. physiol. Psychol.*, 53, 1960, 169-172.

[2] Wodinsky and Bitterman, *op. cit.*, this JOURNAL, 72, 1959, 184-199.

results like those for the rat—initial resistance to extinction was greater after partial than after consistent reinforcement.[3] How are these findings to be understood? One possibility is that partial reinforcement may operate in the same way at the level of the fish as it does at the level of the rat, but that its operation is masked in equated-trials experiments by a greater sensitivity of the fish to frequency of reinforcement. It also is possible, however, to account for existing data on the fish without any reference to a tendency for initial resistance to extinction to be increased by partial reinforcement. The equated-trials outcome may be accounted for in terms of the greater number primary reinforcements given the consistently reinforced animals, while the equated-reinforcements outcome may be accounted for in terms of the greater number of secondary reinforcements given the partially reinforced animals.

One of the purposes served by the research to be reported here was to examine the validity of the secondary-reinforcement interpretation of the equated-reinforcements outcome. In Experiment I, mouthbreeders were *classically* conditioned with a noxious *US;* greater resistance to extinction after partial reinforcement in such an experiment would be rather difficult to explain in terms of secondary reinforcement. In Experiment II, the pattern of reinforced and unreinforced *instrumental* training trials was varied, while the frequency of both kinds of trial remained the same; to account for any differential resistance to extinction appearing under such circumstances, it would be necessary to go beyond the notion of resistance to extinction as a simple function of the frequency of primary and secondary reinforcement. The research to be reported also serves a number of other purposes which will become evident as the separate studies are described.

EXPERIMENT I

In a recent series of equated-trials experiments on classical conditioning in the *goldfish,* no effect of partial reinforcement on resistance to extinction was found; that is, resistance to extinction was neither increased nor decreased by partial reinforcement.[4] The difference between these results and the instrumental results for the mouthbreeder could be attributed either to the difference in species or to the difference in technique. Although the two variables might have been unconfounded in any of a number of designs, it seemed most straightforward to do an equated-reinforcements classical-conditioning experiment with the mouthbreeder. As already noted, such an experiment has important implications for the interpretation of the instrumental data.

---

[3] R. C. Gonzalez, R. M. Eskin, and M. E. Bitterman, Extinction in the fish after partial and consistent reinforcement with number of reinforcements equated, *J. comp. physiol. Psychol.,* 55, 1962, 381-386.

[4] Gonzalez, Longo, and Bitterman, Classical conditioning in the fish: Exploratory studies of partial reinforcement, *J. comp. physiol. Psychol.,* 54, 1961, 452-456; R. C. Gonzalez, Sandra Milstein, and M. E. Bitterman, Classical conditioning in the fish: Further studies of partial reinforcement, this JOURNAL, 75, 1962, 421-428.

Subjects. The *Ss* were 47 sexually mature mouthbreeders about 3 in. long. All previously had been given some training in the instrumental situation, but none had experience either with partial reinforcement or with extinction. The animals were kept in individual 2-gal. tanks on open shelves in a temperature-controlled laboratory.

Apparatus. The apparatus, which has been described in detail elsewhere,[5] was the same as that used in the earlier experiments, with three exceptions. (1) The experimental chambers, enlarged to accommodate somewhat larger *Ss*, were now 5 in. long and 3 in. wide. (2) Both long walls of the chambers were of Plexiglas, one of them (the one near the paddle) having ½-in. holes cut into it at regular intervals. Both walls were painted with a low-resistance mixture of carbon and styrofoam dissolved in methyl-ethyl-keytone through which the shock was administered. (3) The characteristics of the amplifying-integrating circuit were changed to give an increased frequency of response to a standard burst of activity. This change made it possible to rely entirely on counter-readings and to ignore fractional values (previously read from a meter).

Procedure. The *CS* was light-onset, the *CS-US* interval was 4.5 sec., and the offset of the *CS* coincided with the termination of the *US*, a ½-sec. shock of moderate intensity. The *CR* was general activity. It was measured, both in conditioning and in extinction, during the 4.5-sec. interval following the onset of the *CS*.

The *Ss* were divided at random into three groups, a Consistent Group of 15 *Ss*, a Partial-I Group of 16 *Ss*, and a Partial-II Group of 16 *Ss*. On each of the 20 days of training, the Consistent Group had 5 reinforced trials, while the Partial-I Group had 10 trials of which 5 were reinforced (the order of reinforced and unreinforced trials following Gellermann's rules). The training of the Partial-II Group was like that of the Consistent Group for the first 8 days and like that of the Partial-I Group for the remaining 12 days. Then there were 7 days of extinction for all groups during which each animal was given 7 unreinforced trials per day. Throughout the experiment, the intertrial interval averaged 3 min. with a range of 2-4 min.

Results. The principal results of the experiment are shown in Fig. 1. Consistent reinforcement led to a higher level of response in conditioning, but partial reinforcement produced greater resistance to extinction. Both trends are statistically reliable. Analysis of variance based on response to the *CS* in acquisition yields a significant Groups-effect ($F = 6.70$ with

---

[5] J. L. Horner, Nicholas Longo, and M. E. Bitterman, A classical conditioning technique for small aquatic animals, this JOURNAL, 73, 1960, 623-626.

2 and 44 *df*), a significant Days-effect ($F = 42.1$ with 19 and 836 *df*) and a significant Groups × Days interaction ($F = 2.14$ with 38 and 836 *df*). Analysis of variance based on response to the *CS* in extinction yields a significant Groups-effect ($F = 3.66$ with 2 and 44 *df*), and a significant Days-effect ($F = 12.11$ with 6 and 88 *df*), but an insignificant interaction ($F < 1$)—the crossover which is evident in Fig. 1 is not reflected in the interaction, because the analysis of response in extinction excludes Day 20 (the last day of training). In both cases, orthogonal comparisons

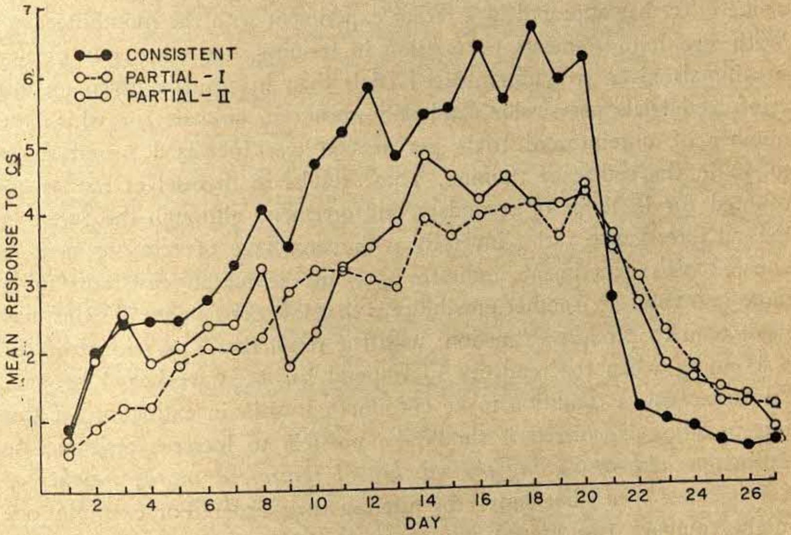


FIG. 1. THE PERFORMANCE OF THE THREE GROUPS IN EXPERIMENT I

show significant differences between the Consistent and the combined Partial groups, but not between the two Partial groups.

These results suggest several conclusions. One has to do with the comparability of the classical and instrumental techniques, at least insofar as the operation of partial reinforcement is concerned: with equated reinforcements, the mouthbreeder shows greater resistance to extinction after partial reinforcement in both types of experiment. The deviant results for the goldfish are of interest in this connection. Not only have classical conditioning experiments failed to show differential resistance to extinction after partial and consistent reinforcement, but so also has a recent instrumental experiment involving controls both for number of trials and for number of reinforcements.[6] Further work with the goldfish obviously is in order. Another conclusion suggested by the present results is that some factor other than secondary reinforcement is responsible for the greater resistance to extinction after partial reinforcement which the mouthbreeder shows when reinforcements are equated. Further evidence on this point is provided by the next experiment.

EXPERIMENT II

When number of primary reinforcements is equated, resistance to extinction in the rat increases with number of unreinforced trials,[7] and a like effect has appeared in a recent experiment with the mouthbreeder.[8] With five reinforcements per session in training, resistance was greater in animals given five unreinforced trials than in animals given no unreinforced trials per session, and still greater in animals for which the number of unreinforced trials per session was increased progressively to 40 in the course of training. These results in themselves can be accounted for in terms of secondary reinforcement, although the fact that the rat's resistance varies inversely with percentage of reinforcement in equated-*trials* experiments indicates that, in the rat at least, something more is involved.[9] Another possibility is that low percentages of reinforcement tend to produce extinction, with the result that *S* is reinforced for responding when the tendency to respond has been weakened by nonreinforcement, a condition to be encountered again in extinction. If this interpretation is correct, it should be possible to increase resistance to extinction *without varying the number of reinforced or the number of unreinforced trials,* but simply by altering their order. For example, one might compare two groups reinforced on 50% of training trials, one with a Gellermann schedule (which does not permit more than three unreinforced trials in succession), and the other with a more extreme schedule (which involves much longer runs of unreinforced trials, and which therefore carries the process of extinction further). Precisely this comparison was made in the second experiment to be reported here.

Another purpose of the experiment was to extend the work on inconsistent reinforcement in the mouthbreeder to the effects of partial delay of reinforcement. Accordingly, two additional groups were included: a consistently reinforced control group (which would serve also as a control for the two partial groups already mentioned), and a group for which reinforcement was delayed 30 sec. on half the training trials. In the rat, partial delay of reinforcement increases resistance to extinction.[10]

---

[6] O. C. Shaffer, Resistance to extinction in the goldfish following partial reinforcement in an instrumental situation, M. A. Thesis, Bryn Mawr College, 1961.

[7] J. M. Notterman, A study of some relations among aperiodic reinforcement, discrimination training, and secondary reinforcement, *J. exp. Psychol.,* 41, 1951, 161-169.

[8] Gonzalez, Eskin, and Bitterman, *op. cit.,* 383-386.

[9] Solomon Weinstock, Resistance to extinction of a running response following partial reinforcement under widely spaced trials, *J. comp. physiol. Psychol.,* 47, 1954, 318-322.

Subjects. The *S*s were 36 experimentally naïve mouthbreeders, about 3 in. long, which were maintained in individual 2-gal. tanks set on open shelves in the temperature-controlled laboratory.

Apparatus. The apparatus employed has been described in detail elsewhere.[11] It consisted of a target which could be lowered into *S*'s living tank, an electronic system for detecting *S*'s contacts with the target, and a pellet-feeder operated by the detecting system. A Standard Electric timer, graduated in units of 0.01-sec., was used to measure latency of response; the clock started with the introduction of the target and was stopped by *S*'s response.

Preliminary training. After adaptation to the experimental situation, the animals were trained to strike at a target baited with Aronson's mixture. Each contact of *S* with the baited target was reinforced with a pellet of food until 10 reinforcements had been earned. This procedure was continued daily until *S* was responding readily (which required three days on the average), and then training with an unbaited target was begun. There were 6 daily sessions with the unbaited target in each of which each *S* earned 20 reinforcements. On the basis of their performance during these six days, the animals were divided into four matched groups of 9 *S*s each: a Consistent Group, a Gellermann Group, an Extended Group, and a Delay Group.

Experimental training. In the first stage of the experiment-proper, there were 30 daily sessions, in each of which each *S* earned 20 reinforcements. The animals of the Consistent Group and of the Delay Group were given 20 trials per day, while the animals of the Gellermann Group and of the Extended Group were given 40 trials per day. For the Consistent Group, each response was reinforced immediately. For the Delay Group, half of each day's responses were reinforced immediately and the remainder were reinforced 30 sec. afterward, the sequence of immediate and delayed reinforcements following selected Gellermann-orders. (On delay-trials, the lever was withdrawn from the water in the interval between response and reinforcement.) For the Gellermann Group, 20 (immediately) reinforced trials and 20 unreinforced trials were given each day, which was true also of the Extended Group. For the Gellermann Group, of course, the two kinds of trial were scheduled according to Gellermann's rules throughout the training. For the Extended Group, Gellermann's rules were used on the first two days only; then the order was changed to involve progressively longer runs of unreinforced trials up to a maximum of 17 on Day 13. Thereafter, there was a run of 17 unreinforced trials each day, its position in each day's series of trials (early, middle, or late) being varied systematically in counterbalanced fashion. Trials were massed for all groups; the intertrial interval averaged about 5 sec.

---

[10] Janet Crum, W. L. Brown, and M. E. Bitterman, The effect of partial and delayed reinforcement on resistance to extinction, this JOURNAL, 64, 1951, 228-237; F. A. Logan, E. M. Beier, and W. D. Kincaid, Extinction following partial and varied reinforcement, *J. exp. Psychol.*, 52, 1956, 65-70.

[11] Longo and Bitterman, Improved apparatus for the study of learning in fish, this JOURNAL, 72, 1959, 616-620.

On Day 31 (which was not, of course, the same calendar day for all *Ss*—their training was staggered), all *Ss* were extinguished. The extinction-trials were the same in every respect as the unreinforced training trials. When, on any trial, an animal failed to respond in 30 sec., the target was withdrawn, the time was recorded, and the target was reinserted in the usual way to begin the next trial. The criterion of extinction was five successive failures to respond in 30 sec.

Results. Mean log latency of response in all groups declined in negatively accelerated fashion to a stable level (of about 0.25 sec.), which was reached by the twentieth day of training. Latency of response in the Delay Group tended to be higher than that in the other three groups, but the differences among the four groups were not statistically significant ($F = 1.6$ with 3 and 28 *df*), nor was the interaction of Groups and Days ($F < 1$). The variance due to days was significant at the 1% level ($F = 5.5$ with 29 and 812 *df*). The analysis is based on the data for 9 Consistent, 8 Delay, 8 Gellermann, and 7 Extended *Ss*, 4 animals having been lost for one reason or another in the course of training.

Early in training, both partial groups (Gellermann and Extended) showed the tendency—characteristic of partially reinforced mouthbreeders—to respond more rapidly after unreinforced trials than after reinforced trials. As training continued, the tendency was reversed in the Extended Group, and the change could be traced to the effects of the long runs of unreinforced trials, in the course of which the Extended animals began to extinguish. Since there was almost immediate recovery with the reintroduction of reinforcement at the end of these runs, the over-all mean latency of the Extended Group was not very different from that of the others.

The course of extinction in the first 70 trials is plotted in Fig. 2. A mean log time for each block of five trials was computed for each animal, and the median of those means was used for the plot. The mean of the means gives much the same picture as the median, except that it does not show quite as dramatically the abrupt shift in the performance of the Gellermann animals which occurred after about 40 trials. (The same sort of abrupt transition appeared in the data of an earlier matched-trials experiment,[12] Gellermann and Extended curves running along together for many trials and then diverging sharply.) Analysis of variance based on mean log time per block of five trials yields a significant Groups-effect ($F = 3.84$ with 3 and 28 df), a significant Days-effect ($F = 45.58$ with 13 and 364 df), and a significant interaction ($F = 1.77$ with 39 and 364 df). The Extended Group differed significantly from the other three ($F = 9.36$ with 1 and 28 df), but Gellermann did not differ significantly from Consistent and Delay, nor did Delay differ significantly from Consistent. A simple analysis of trials to criterion also showed a significant Groups-effect ($F = 5.31$ with 3 and 28 df) which reflected primarily a significant difference between the Extended Group and the other three ($F = 12.7$ with 1and 28 df).

---

[12] Gonzalez, Eskin, and Bitterman, *op. cit.,* 383-386.

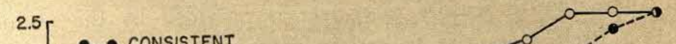


FIG. 2. THE COURSE OF EXTINCTION IN EXPERIMENT II

These results, taken together with those of Experiment I, show clearly that the mouthbreeder's resistance to extinction is increased by inconsistent reinforcement in training. Both sets of results can be accounted for on the assumption that unreinforced trials have certain discriminable consequences which (as in Experiment I) may enter into the *CS*-complex or (as in Experiment II) play a role in the elicitation of instrumental response. It may be that the Hull-Sheffield interpretation, in terms of sensory carry-over,[13] will do nicely for the fish, however inadequate it may be for the rat. The failure of partial delay of reinforcement in Experiment II to increase resistance to extinction as it has in previous experiments with the rat provides some support for this interpretation; the poor showing of the Gellermann Group weakens the argument to a certain extent, although the shape of the Gellermann curve does suggest that partial reinforcement had an effect which partial delay did not. It would be interesting to look at the results for an Extended Delay Group —that is, a group trained with longer runs of trials on which reinforcement was delayed.

---

[13] V. F. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.*, 39, 1949, 511-526; C. L. Hull, *A Behavior System*, 1952, 120-121.

An alternative to sensory carry-over is provided by the notion that extinction itself gives rise to certain sensory events which the animal may discriminate. Thus, reinforcing a partially extinguished fish in the instrumental situation may serve to encourage response under the sensory conditions which characterize partial extinction. From this point of view, the difference between the Gellermann and Extended Groups of Experiment II may be due to the greater amount of extinction produced in the Extended animals by the long runs of unreinforced trials. The fact that the resistance of the Gellermann group relative to the consistently reinforced control was less here than in a previous equated-reinforcements experiment, in which there were only 5 reinforcements per day,[14] may mean that 3 unreinforced trials in succession produce less extinction in the context of 20 reinforcements per day than they do in the context of 5. An experiment with widely spaced training trials would readily distinguish this interpretation from the interpretation in terms of sensory carry-over; spacing should prevent sensory but not (what might be called) extinctive carry-over.

To what conclusions are we led at the present time by the work on partial reinforcement in the fish? The earliest experiments—with equated trials—seemed to show fish and rat to be differently affected by partial reinforcement, but in the light of subsequent work with equated reinforcements the first results were seen to be a function only of differential sensitivity to frequency of reinforcement. Increased initial resistance to extinction in the fish after partial reinforcement has not yet been demonstrated with equated trials, but there is little reason to doubt that it could be, if the schedule were a properly extended one. Given the equated-trials results of Wodinsky and Bitterman for a series of extinctions separated by retraining sessions—lower asymptotic resistance in the consistently reinforced animals[15]—the experiment hardly seems worth the effort.

---

[14] Gonzalez, Eskin, and Bitterman, *op. cit.*, 383-376.

[15] Wodinsky and Bitterman, *op. cit.*, this JOURNAL, 72, 1959, 184-199.

What other differences remain to be considered? First, there is the fact that partial delay of reinforcement has not produced in the fish the increased resistance to extinction that it does in the rat. Secondly, there is the fact (not yet mentioned here) that alternating and Gellermann schedules do not produce in the fish the differential resistance to extinction that they do in the rat.[16] Finally, there is the suggestion that the Gellermann schedule is not as effective for the fish as it is for the rat, either from the standpoint of initial resistance to extinction (note the lack of a significant difference between the Gellermann and Consistent Groups of Experiment II), or from the standpoint of asymptotic resistance (note the rapid disappearance of the initial difference, found in a previous experiment,[17] in the course of a series of extinctions separated by retraining sessions). It is difficult to know how much weight should be given these essentially null outcomes, but they cannot be ignored. As matters now stand, *all of the data for fish can be accounted for in terms of a simple carry-over concept (either sensory or extinctive) which clearly is inadequate at the level of the rat.*

SUMMARY

Two experiments are reported on partial reinforcement in the fish. In the first, a classical conditioning technique was used, with shock as the *US*. With number of reinforcements equated, partial reinforcement produced greater resistance to extinction than did consistent reinforcement. In the second experiment, an instrumental technique was employed, the animals being trained to strike at a target for food-reward. No difference in resistance to extinction was found between a consistently reinforced group and a group for which reinforcement was delayed on half the training trials. A partially reinforced group for which the schedule of reinforcements and nonreinforcements followed Gellermann's rules showed much less resistance to extinction than did a group with the same number of reinforced and unreinforced trials scheduled in such a way as to involve much longer runs of unreinforced trials. The Gellermann animals showed some tendency to resist extinction longer than the consistently reinforced animals, but the difference was not statistically significant. The implications of these findings are considered.

---

[16] Gonzalez, Eskin, and Bitterman, Alternating and random partial reinforcement in the fish with some observations on asymptotic resistance to extinction, this JOURNAL, 74, 1961, 561-568; D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, this JOURNAL, 66, 1953, 57-65.

[17] Gonzalez, Eskin, and Bitterman, *op. cit.*, *J. comp. physiol. Psychol.*, 55, 1962, 384.

DECISION-TIME WITHOUT REACTION-TIME: EXPERIMENTS IN VISUAL SCANNING

By ULRIC NEISSER, Brandeis University

It has become common to regard human beings as processors of information, but not much is known about the underlying functions. Accurate measurement of the *time required* to execute cognitive operations may help us to interpret them. In the classical reaction-time experiment, processing-time is confounded with the time used to prepare and execute the physical response itself. In the present experiments, a procedure involving visual scanning was used to circumvent this problem. An *S* who scans a list of items, to find one of a certain kind, must examine each item he encounters, but he makes no response until he comes upon the particular item for which he is searching. The rate at which he scans measures the time he uses to analyze the items that elicit no response.

The five experiments reported here serve two purposes. First, they 'calibrate' the scanning method, which can be considered trustworthy to the extent that different *Ss* in different experiments produce comparable results. Secondly, they provided preliminary information about the depth, breadth, and flexibility of the processes involved in recognizing printed letters.

The experiments are designed and interpreted on the assumption that the process of recognition is hierarchically organized. Before an *S* 'decides' that the letter *Z*, for example, is present in the input, he must make prior 'decisions' about subordinate features such as parallel lines and angles; these in turn are probably based on processes of a still lower order. We should expect processing-times to depend on the depth of the hierarchy required by a problem. If, however, several operations are at the same level in the hierarchy, *S* may be able to execute them simultaneously[1] The scanning method may enable us to determine whether simultaneity actually is possible. It may also provide information about

---

* Received for publication December 4, 1961. This research was supported by the Lincoln Laboratory of the Massachusetts Institute of Technology, operated with support from the U.S. Army, Navy, and Air Force.

[1] The theoretical significance of such simultaneity is explored in Ulric Neisser, The multiplicity of thought, *Brit. J. Psychol.*, 54, 1963, 1-14.

the flexibility with which *S* can shift from one mode of processing to another, to take advantage of altered circumstances.

Method. A list of 50 items arranged in a singe-spaced column is presented to *S*. Each item is a string of letters. As soon as the list is exposed, a clock starts, and *S* begins to scan down the list from the top. When he finds the single item that has the *critical property,* he turns a switch which stops the clock. If the critical property is the absence of *Z*, for example, the list might contain 49 items like *JZTXVB, DQFJHZ, MBZJSV, ZXLSMT, RLZQXS,* and one critical item like *VXRLFH*. A series of 15 such trials (lists) with the same critical property usually is given; the critical item appears in an unpredictably different position on each trial. Afterwards, scanning-time is plotted as a function of the position of the critical item, *i.e.* of the number of items scanned. These graphs usually are fitted fairly well by straight lines. An example appears in Fig. 1. Linearity implies that the time taken to scan each item does not change from one end of the list to the other. The most important property of such a line is its *slope:* the average time per item scanned. The slope represents the time *S* needs to assure himself that each item does *not* have the critical property. In the example above, he must identify a *Z* in each item.

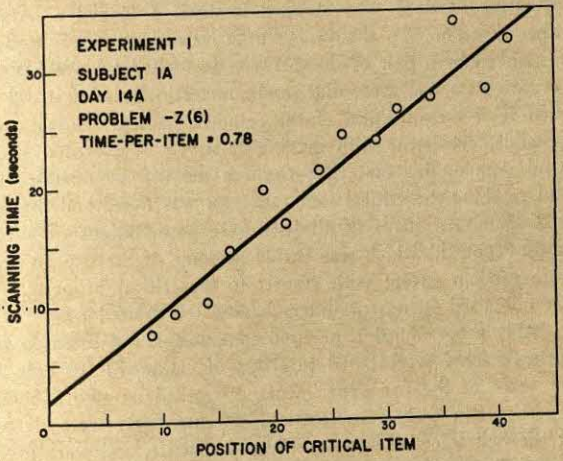


FIG. 1. SCANNING-TIME AS A FUNCTION OF THE POSITION OF THE CRITICAL ITEM: A TYPICAL GRAPH

The slope is unaffected by the time required to begin scanning, to decide upon a response, or to turn the switch. There is no reason to believe that any of these response-factors varies with the position of the critical item on the list. They affect only the intercept of the fitted line—its height above the *X*-axis.[2] The slope, or time-per-item, therefore is a relatively pure measure of the time required to process the information. If *S* scans at the fastest rate consistent with relatively error-free performance, his rate (in the example) should be limited only by the speed with which he can analyze the items for the presence of *Z*. (Physiological limitations on eye-movements may limit the speed to some maximum, but would not affect the conclusions drawn here.) In some problems (*i.e.* with some types of critical property), *S* may process the items in groups rather than individually. In this case 'time-per-item' is a fiction, but a convenient one; it still enables us to compare 'processing-time' for different tasks.

---

[2] Some non-linearity is to be expected at the beginning of the list, as *S* begins to scan. For this reason, the *y*-intercepts of the fitted lines do not provide reliable information.

The *S*s were college students, paid for their services. The lists were exposed in a device which activated a Standard Electric Timer (calibrated in 0.01 sec.) when *E* opened a spring-loaded door, and stopped it when *S* turned a hand-held rotary switch in either direction. To discourage premature responses, alternate items on all lists were flanked by a pair of dots; *S* was instructed to turn his switch clockwise if the critical item had these dots, and counterclockwise if it did not. (Signal lights informed *E* of the direction of the response chosen.) *S* was urged to scan as fast as possible, consistent with making relatively few errors. If he failed to find the critical item on his first time through the list, or mistakenly turned the switch without reaching the critical item, an error was noted and the same list was later presented again. All items in all lists were permutations of a certain set of letters called the 'context,' which was varied in some of the experiments. The permutations were random except with respect to the critical property. All the lists, made up of randomized strings of letters having the desired properties, were prepared by a suitably programmed IBM 7090 computer, and printed on an ANELEX high-speed printer. They were taped to strips of cardboard for ease in handling.

In the first three of the five experiments, *S* worked on two different problems (critical properties) in each daily session. Twenty trials were devoted to each problem, but the first six trials were considered practice, and were not used in the determination of slope. A daily session took from 30-45 min. The critical items actually occurred at list-positions 5, 6, 25, 30, 45, and 46 (randomly ordered) during practice, and at 9, 11, 14, 16, . . . , 39, 41 (randomly ordered) during the experiment proper, but *S* did not know this. In Experiments 4 and 5, only 15 trials were given for each problem, the first three being practice. The position of the critical item in each trial was randomly chosen from the integers 1, 2, . . . , 50, except that it was never the same twice for any one problem on one day. In these experiments, *S* worked on 4-5 problems in each daily session.

In every case, the points were plotted and a straight line was visually fitted to them. The data reported here, however, are not taken from these visual fits but from lines fitted mathematically by the method of least squares. Slopes and intercepts of the best-fit lines, as well as standard errors of estimate around the lines, were calculated by an IBM 7090 computer. The computer-program disregarded any isolated points which deviated greatly from the line determined by the others. (Actually, the program discarded points which deviated from the line at 5% level by a *t*-test. This criterion was used for the sake of simplicity; it is not fully satisfactory because search-times are markedly more variable when the critical item is near the bottom of the list than when it is near the top.)

EXPERIMENT I: PRELIMINARY

The first experiment was a preliminary study of the effects of the following variables:[3] the identity of the target-letter (*Q* vs. *Z*); the number of letters in each item (2 vs. 6); the complexity of processing required (scanning items *without* the target-letter to find an item that contains it versus scanning items *containing* a letter to find an item without it); the number of acceptable alternative target-letters (1 vs. 2).

Procedure. The design of the experiment involved 14 conditions: 7 critical properties, each embedded in two sets of lists. In one set, each item was two letters wide (*e.g. ZD, JZ, LR, ZQ*), while in the other set each was six letters wide (*e.g. JZLXSH, QVZMXL, FDRVQH*). The critical properties were *Z, Q, ZvQ, —Z, —Q, —ZvQ,* and *—QvZ*. These terms are to be interpreted as follows. In Problem *Z,* the critical item alone contains a *Z,* which may be considered the target of the search; in *ZvQ* the critical item may contain either *Z* or *Q* or both, and *S* does not know which until he encounters it (parallel scanning); in *—Z,* all the items *except* the critical one contain a *Z*; in a problem like *—ZvQ*, all items except the critical one have a *Z* and no *Q*, while the critical one may lack a *Z,* may have a *Q*, or both. The lists were made up of *J, P, S, T, V,* and *X* in addition to *Z* and *Q*.

Three *S*s were used.[4] Each was given 2 practice sessions, 7 experimental sessions in which the 14 problems were presented (in a different order to each *S*), and 7 additional sessions in which the 14 problems were presented again in the opposite order.

Results. Since data for the different *S*s are very similar, only mean times-per-item appear in Table I. We may draw certain tentative conclusions. (1) *–Z* takes longer than *Z*, and *–Q* takes longer than *Q;* (2) *QvZ* takes no longer than *Z* alone; (3) six-letter lists take longer than two-letter lists, but generally not three times as long; (4) *Z* takes longer than *Q*, and *–Z* takes longer than *–Q;* (5) in the two-letter lists, all problems involving the absence of a letter take about equally long. While many of these findings will be extended or confirmed in the four experiments to follow, we will pause at this point to consider their implications for the depth, breadth, and flexibility of the cognitive processes involved.

---

[3] Experiment 1 has been described earlier: Neisser, Time-analysis of logical processes in man, *Proc. West, Joint Computer Conf.*, 1961, 579-585). Since the results presented in that paper are based on visually fitted lines, they differ slightly from those appearing here.

[4] The assistance of Miss Emily Carota, Mr. Arthur Warmoth, and Mr. Norman Goldberg, who served as *E* in Experiments I, II, and III, of Mr. Robert Greenway, who served as *E* in Experiments IV and V; and of Mr. Paul Weene, who assisted in the construction of the apparatus is gratefully acknowledged. Experiments III and V were carried out at the Harvard Center for Cognitive Studies.

When *S* is looking for a particular letter (*e.g. Z*), each noncritical item gets only scant attention. *S* need only view each item long enough that the lower-order recognitive systems could (for the individual features of *Z*) be activated by the right input. That input does not appear, and the *Z*-recognizer as a whole is not activated, until the critical item is reached. This analysis is substantiated by introspective report: *S* does not 'see' the letters he passes; he 'sees' only a blur until the Z 'jumps out at him.' The situation is different when *S* looks for an item without a *Z*. Now the *Z* in each item must be identified; the full depth of the recognizer for *Z* is used on each line. Because of this greater depth, problems in which *S* looks for the absence of a letter take longer than those in which he need only find the letter. The effect of practice in this difference will be examined in Experiment IV.

Even in the rapid search for a target, *Z*, elementary analyses of the stimulus-input are being carried out. The elementary operations that could detect a *Z*, if it were to occur, must be different from those which could detect *Q*. Since *S*s can be alert for *both* letters without slowing their scanning-rate, it seems that the different operations can be carried out simultaneously. (It may be argued that the operations for *Q* are somehow interleaved with those for *Z*, instead of being simultaneous with them, a point which will be considered further in the light of the data of Experiment V.) This evidence for the breadth of the processes involved is in contrast to a finding which seems to emphasize the opposite: wider lists take more time to scan. One might easily imagine a six-channel device able to handle two- and six-letter widths with equal speed, but no *S* behaves like such a device. In Experiment II, we shall consider whether this effect is due to the defects of peripheral vision; in Experiment IV, whether it disappears with practice.

TABLE I

MEAN TIMES-PER-ITEM (SEC.) IN EXPERIMENT I
(Each entry is the mean of the slopes of the 3 *S*s, each slope based on about 14 measurements.)

No. of letters	Day	Problem						
		Q	*−Q*	*Z*	*−Z*	*QvZ*	*−QvZ*	*Qv−Z*
6	1	0.14	0.42	0.50	0.70	0.60	0.98	0.82
	2	0.08	0.42	0.55	0.64	0.45	0.72	0.65
2	1	0.09	0.32	0.23	0.37	0.21	0.30	0.34
	2	0.05	0.27	0.20	0.29	0.20	0.31	0.34

The difference in speed between *Z* and *Q* may be an unchangeable property of the recognitive systems for these two letters. If, however, these systems are flexibly organized, we should expect the difference to be eliminated or reversed in a different letter-context or a different experimental situation. The first of these predictions is tested in Experiment III. The second is borne out, rather unexpectedly, by the observation that *–Z*, *–Q*, *–ZvQ*, and *–QvZ* are equally fast in two-letter lists. A glance at the lists themselves provides an explanation. In these problems, there is a *Z* in every non-critical item; the result is a *column* of *Z*s which switches haphazardly from one side of the list to the other. Following such a column with the eye is very different from inspecting individual items. Evidently the *S*s discovered this, and promptly used it to increase the speed of their scan. No such columns are formed in six-letter lists.

EXPERIMENT II: HORIZONTAL SPACING

Two reasons might be advanced for the fact that it takes less time to process a two-letter item than a six-letter item. First, the six-letter item extends farther horizontally, and thus may encourage real or incipient eye-movement, or may require parafoveal vision. (The six-letter item occupies about 5° of visual angle.) Secondly, the increase in amount of information to be processed may be the important variable. To decide this issue, horizontal spacing itself was varied in Experiment II.

Procedure. There were four conditions. In each case the target was *Z*, and the letter-context was identical with that of Experiment I. The conditions were *Z(2)* (items two letters wide as in Experiment I); *Z(4)* (items four letters wide); *Z(4w)* (items composed of four letters, but with two inserted dashes to increase the width of the item, as *P-TX-Q*); *Z(6)* (items six letters wide, as in Experiment I). Each of 4 *S*s scanned in each condition five times (two per day).

TABLE II

MEAN TIME-PER-ITEM (SEC.) IN EXPERIMENT II
(Each entry is the mean of the slopes of the 4 *S*s, each slope based on about 14 measurements.)

Day	Problem			
	Z(2)	*Z*(4)	*Z*(4*w*)	*Z*(6)
1	0.22	0.41	0.48	0.79
2	0.20	0.30	0.30	0.51
3	0.17	0.27	0.25	0.34
4	0.14	0.20	0.21	0.29
5	0.11	0.16	0.16	0.25

Results. Mean times-per-item appear in Table II. It is evident that Condition *Z(4w)* is just as fast a Condition *Z(4)*, and systematically faster than *Z(6)*. In other words, the critical factor is the number of letters per item, rather than their horizontal spread. Note that, on the first two trials, times-per-item for *Z(2)* and *Z(6)* are comparable with those of Experiment I, but that practice causes sharp decreases in time. The effects of practice will be considered further in Experiments IV and V.

EXPERIMENT III: CONTEXT

The difference in time-per-item between problems involving *Z* and those involving *Q* was one of the most striking results of Experiment I.

The difference might be ascribed to the two letters as such or to the letter-contexts in which they were embedded. The other letters in Experiment I were *J, P, S, T, V,* and *X*. Perceptually, the *Q* seems to stand out vividly from this assortment because of its roundness. For Experiment III, two letter-contexts were selected. One was intended to make *Z* even harder to see, the other to obscure the *Q*. The first, called the 'angular' context, consisted of the letters *E, I, M, V, W,* and *X;* the second, called the 'round' context, consisted of *C, D, G, O, R,* and *U*. Six-letter items were used.

Procedure. The 3 *S*s worked on six different problems, doing each problem twice in each context. In addition to studying the context-effect, we wished to confirm and extend the finding of Experiment I that it takes no longer to scan for two letters in parallel than for one alone. The problems used were, therefore, *Q, Z, QvZ, –Q, –Z,* and *–Qv–Z*. The last of these represents a list in which every item except the critical one has both a Q and a Z (in either order, not necessarily adjacent), while one or both of these letters are absent from the critical item. In other respects the design was like that of Experiment I.

TABLE III

MEAN TIMES-PER-ITEM (SEC.) IN EXPERIMENT III
(Each entry is the mean of the slopes of the 3 *S*s, each slope based on about 14 measurements.)

Letter-Context	Day	Problem					
		Q	*–Q*	*Z*	*–Z*	*QvZ*	*–Qv–Z*
Angular	1	0.08	0.36	0.24	0.45	0.13	0.74
	2	0.07	0.35	0.22	0.49	0.17	0.61
Round	1	0.58	0.68	0.11	0.34	0.60	0.82
	2	0.46	0.61	0.09	0.32	0.56	0.77

Results. Mean times-per-item appear in Table III. It is evident that contextual background is of great importance in determining processing-time. In the angular context, problems involving *Q* take much less time than those involving *Z,* while the opposite is true in the round context. Indeed, the *S*s found the search for *Q* or for *–Q* very difficult against the round background, and made frequent errors based on confusions between *Q* and *O*, or *Q* and *G*.

The results confirm Experiment I in showing that it takes no longer per item to scan in parallel for *Q* or *Z* than for one of these targets alone: no longer than to scan for *Q* in the round context, and no longer than *Z* in the angular context. No such parallelism is demonstrable, however, at the more complicated level represented by *–Q, –Z,* and *–Qv–Z*. Here *S* must assume himself that each item he scans *does* contain the critical letter, or letters. To check for both takes consistently longer than to check for one alone.

EXPERIMENT IV: EXTENDED PRACTICE

The results of Experiment II demonstrated that scanning time decreases with practice, at least with a simple problem like *Z*. It is particularly important to determine whether the relation between the time-per-item for different types of problems is also altered. Experiments I and II showed that wider items take longer to analyze; will this remain true after extended practice? Experiments I and III showed that problems involving letter-absence, which require *S* to identify the critical letter on each line, take longer than problems in which a single instance of a letter is to be located. Will this also remain true after extended practice?

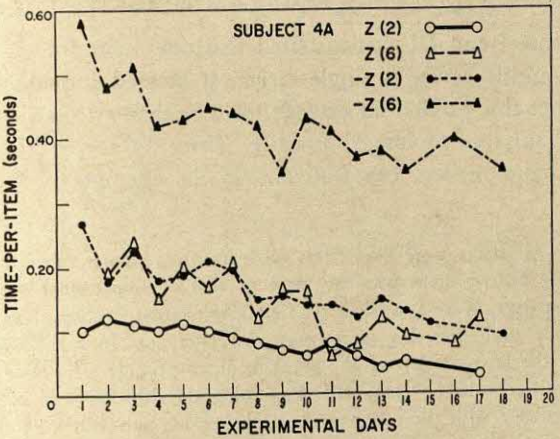


FIG. 2. TIME-PER-ITEM AS A FUNCTION OF DAY OF PRACTICE
Each point represents a single slope, based on about 12 measurements.

Procedure. Each of 3 *Ss* was given 17 or 18 sessions on more or less consecutive weekdays. Four problems were used: *Z* and *–Z*, each in two-letter and six-letter widths. The letter-context included *B, D, F, H, J, L, M, Q, R, S, T, V*, and *X* as well as *Z*. Usually, all four problems were presented each day, with the order of presentation varied from one day to the next.

Results. The results for one *S*, plotted in Fig. 2, show that items six letters wide take at least twice as long as those two letters wide, even after three weeks, and that *–Z* takes much longer than *Z* in items of both widths. The results for the other two *Ss* (which cannot be presented for lack of space) substantiate these conclusions. There were, however, pronounced individual differences in the effect of practice. One of the other two *S*s began with times-per-item from 50-200% longer than those of the *S*, whose data are plotted, and much larger day-to-day variation. The third *S* began more slowly still, but improved dramatically, and attained speeds comparable to those of the first *S* by the third week. These differences are comparable to those stressed by Bryan and Harter in their classical study of telegraphy as a congnitive skill.[5]

EXPERIMENT V: EXTENDED PRACTICE WITH PARALLEL SCANNING

Experiments I and III demonstrated that one scans for either of two targets as quickly as for a single target. It seemed important to determine whether this parallelism extends to more than two targets, and also whether it survives prolonged practice. Three *S*s served for 30 or 31 weekdays, more or less consecutively, in an attempt to answer these questions.

Procedure. All items were six letters wide (with a minor exception noted below), and the letter-context was the same as that of Experiment 4. On each of the first seven days, every *S* worked on four single-target problems: *H, M, Q,* and *Z*. These four problems also were presented later, on four different days interspersed in the later portion of the experiment, to provide a baseline with which parallel scanning could be compared. Beginning on the eighth day, each *S* was presented regularly with three two-target problems instead: *HvM*, *HvZ*, and *QvZ*. On the thirteenth day, a four-target problem, *HvMvQvZ,* was introduced and became a regular part of the procedure. In this problem, the critical item might contain *any* of the four target letters, and *S* did not know which until he found it. Of course, no non-critical item contained any of the four letters. On the last few days of the experiment (Day 26 or later), another problem was introduced for exploratory purposes. This was *HvMvQvZ* in a format *eight* letters wide instead of six. After 4-6 days of practice with this problem, none of the *S*s has achieved the speed with which they were doing *HvMvQvZ (6)*. As far as it goes, this finding seems to confirm the results of Experiment IV, but we will not consider it further here.

Results. The results for one of the *S*s are plotted in Fig. 3. The other two *S*s produced very similar data. It is evident that parallel scanning can be sustained and that it applies to situations involving four targets. All three *S*s were scanning the two- and four-target problems with about equal speed in the last week. This speed was generally within the range of speeds used on the single-target problems, although slightly above the mean of the speeds on the four single targets. It appears consistently possible to be on the lookout for several things at once.

---

[5] W. C. Bryan and Noble Harter, Studies in the physiology and psychology of the telegraphic language, *Psychol. Rev.* 4, 1897, 27-53.

It would be difficult to maintain that these results are due to interleaving. There might indeed be enough dead time between analyses for slanting lines, say, to carry out the detection of roundness, but one would expect these dead times to diminish with practice; *Z* should gain more from practice than *ZvQ*. The data contradict this prediction, because times-per-item for the different problems converge with practice instead of separating. The finding with *HvMvQvZ* also argues for simultaneity rather than for interleaving. How much dead time can there be?

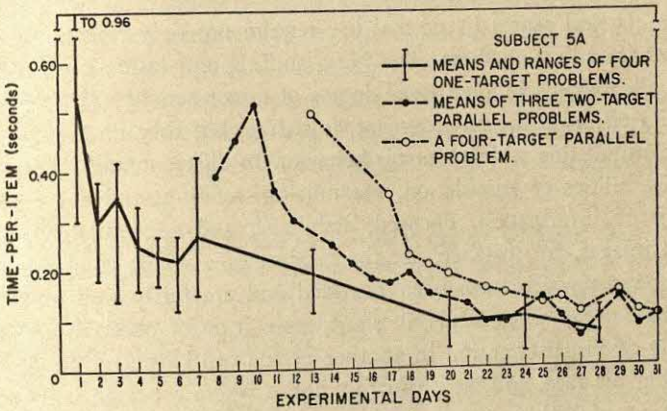


FIG. 3. TIME-PER-ITEM AS A FUNCTION OF DAY OF PRACTICE
For the one-target problems, each point is based on four slopes; for the two-target problems each is based on three; there is a single four-target problem. Each slope is based on about 12 measurements.

SUMMARY

The method of visual scanning was employed in five experiments to obtain direct measures of the processing time of human information. The results indicate that the method is reliable and permit several tentative conclusions about the organization of cognitive processes in the identification of printed letters. (1) At simple levels, several distinct processes of recognition can function simultaneously in the analysis of a single stimulus-configuration. (2) No such simultaneity appears in the analysis of *spatially distinct* parts of the input, even after extended practice. (3) The recognitive hierarchy for a given task can be altered to take advantage of different contextual or other conditions. (4) Positive identification of a letter (such as is necessary when a response is contingent on its absence) takes longer than the simple search made when the response is to be contingent on its presence.

PERCEPTION OF ANOTHER PERSON'S LOOKING BEHAVIOR

By JAMES J. GIBSON and ANNE D. PICK,
Cornell University

The classical senses in normal use require not only receptors but also muscles for adjusting them. Touching, sniffing and tasting, listening, and above all, looking, involve some degree of motor activity. These responses have no effect on the environment, it is true, but they do have an effect on the stimulation which controls behavior. In different ways, they explore the possibilities of stimulation, maximizing, refining, and sharpening its content of information. Focused, stabilized, and appropriately centered retinal images, for instance, are the product of a whole complex of ocular responses which have been recorded and are fairly well understood. The sense-organ adjustments, in short, are a form of observable behavior. The act of visual attention in another person can be perceived simply by watching his face, and we do in fact notice where a person is looking.

The act of looking can be treated as a source of stimulation as well as a type of response. The eyes not only look, but are looked at. The direction of a person's gaze usually indicates what object he is interested in or what person he is responding to in the sphere of the environment, and a shift of his fixation may indicate what he is going to do next, as the boxer and the fencer know. The fact of *being looked at,* that is, a line of gaze fixed on the body of the observer himself, is especially significant.

That men do discriminate, or try to discriminate, the postures and movements of the eyes of another person is strongly suggested by the number of words used in ordinary speech to distinguish modes of looking. Variants of the verb 'to look' include *stare, watch, peer, glance, peep, glare, glower, gaze, contemplate,* and *scan.* Adverbial modifiers of the verb are frequent. A person can be said to look *directly,* or *askance, overtly* or *covertly, boldly* or *bashfully, sternly* or *mildly, critically* or *kindly,* or even *unseeingly.* Among the many idiomatic expressions or metaphors are, *he caught my eye,* or *held my eye,* or *looked me up and down,* or *his eyes dropped,* or *flickered,* or *his gaze wandered.* A person may either *cast an eye on,* or *fasten his eyes on,* or *look down his nose at.* He may *steal a glance,* give a *sidelong glance,* or a *guarded glance,* or have a *discerning glance,* a *piercing glance,* or a *fixed stare.*

---

* Received for publication January 27, 1962. This research was supported in part by the Office of Naval Research under Contract 401 (14) with Cornell University. Special thanks are due to Jane Danielson for serving as the 'looker' in this experiment. The work has benefited greatly from a previous unpublished study by Elinor Wardwell.

He may also give one a *sly look,* an *open look,* or a *black look.* These examples suggest that the task of judging expressions of emotion entails a separate and distinct problem, that of judging what might be called the expression of *attention.*

When we observe that one person looks at another angrily (or sadly, fearfully, or happily), we are observing a space as well as an emotion. A face is normally seen surrounded by an environment in three dimensions. The instigator of emotion is visible as well as the face. A portrait, therefore, can render an emotion but not the object or event provoking it, as evidenced by the direction of the line of regard.

The facial patterns for the perception of emotion have been studied in many experiments, but the ocular patterns for the perception of the act of looking at something have been neglected. Only a few studies are even relevant.

Cline investigated the perception aroused by a pair of schematic faces in three-quarters profile whose eyes were so drawn as to appear to be looking at each other.[1] He systematically varied the expressive features of the two drawings and noted the social situations perceived in accordance with these combinations. But he did not vary the apparent gaze-line of either face.

Wardwell, assuming that being looked at is a potent social stimulus, studied the effect on the behavior of a child of an adult who stared at him most of the time, part of the time, or none of the time.[2] In a test-situation, although the children were preoccupied with the task, they proved to be very conscious of whether they were or were not being observed, and for what proportion of the time.

Recently, Wada has noted that rhesus monkeys are aware of being looked at by a human individual, the criterion being a specific change in electrical activity recorded from the brain-stem.[3] Cornell rhesus monkeys also seem to behave differently when a pair of eyes is fixed on them than when the same pair of eyes is fixed away from them, as determined by an invisible observer looking through a one-way screen. Perhaps all primates perceive whether or not they are being looked at. If so, they must have acute discrimination for the line of gaze of another individual.

Stimulus for the perception of the act of looking. The potential stimulus for the perception of the act of looking must be distinguished from the *source* of the stimulus, although experimenters often do not do so. The stimulus proper is the light reflected from the 'looker,' more exactly, the differential rays of light which can enter the eyes of the *O*. The source of stimulation is the head and eyes of the looker in their surroundings. The posture of the eyes relative to the head and the surroundings is what has to be perceived by *O*. It can only be given in the bidimensional array of light reflected by these surfaces. The face itself is an elastic mobile surface of skin, and the visible eyes are the bulging frontal surfaces of spheres. The face has its 'features,' and the eyes have a pupil and sclera within a frame formed by the eyelids. These shapes are projected to *O* as an arrangement of forms within larger forms.

---

[1] M. G. Cline, Some perceived properties of social interaction. Unpublished Doctoral dissertation, Cornell University, 1954.

[2] Elinor Wardwell, Children's reactions to being watched during success and failure. Unpublished Doctoral dissertation, Cornell University, 1960.

[3] J. A. Wada, Modification of cortically induced responses in brain stem of shift of attention in monkeys, *Science*, 133, 1961, 40-42.

As a first approximation, one might suppose that the stimulus-information for the perception of the looker's gaze is a form-variable connected with 'centering,' that is, the concentricity of the pupil in the sclera, the 'white of the eye.' When it is centered, the gaze will be seen to be directed at *O*; when it is off-centered in either direction, the gaze is perceived correspondingly to deviate toward the right or left of *O*. Fig. 1 isolates this variable and illustrates the hypothesis. One might in fact study experimentally the acuity with which iris-centering and off-centering are detected. This approximate hypothesis, however, is insufficient. On further consideration, one realizes that it holds for the special case when the looker's head is pointing in the same direction as his eyes, in the frontal or facing position. The fact is that a person does not always face in the same direction as he looks. A face can be recognized in a series of stimulus-transformations varying from full-face to profile. Hence this form-variable, the face-form, must also be taken into account in analyzing the stimulus-information. Fig. 2 illustrates the fact that the stimulus-meaning of the eye-forms depends on the face-form that surrounds it. The illustration brings up the whole question of the influence of one form on

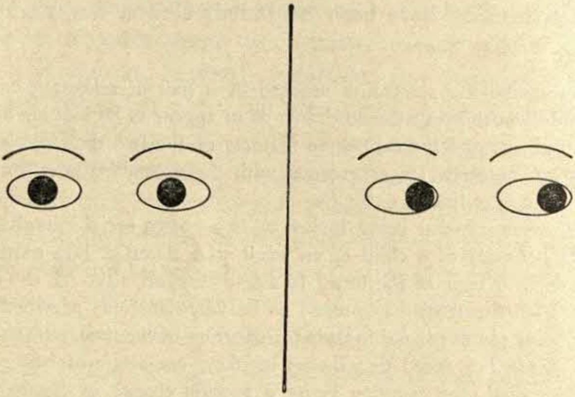


FIG. 1. SCHEMATIC CENTERING OR OFF-CENTERING OF THE PUPIL IN THE SCLERA AS THE STIMULUS FOR THE IMPRESSION OF BEING OR NOT BEING LOOKED AT

another included within it; we need only note here a quite simple relation —that the asymmetry of the projected face is equal and opposite to that of the projected eyes when the eyes are perceived to be looking at 'me.'

Physiologically, a turning of the eyes normally compensates for the opposite turning of the head, this being the reflex mechanism by which the eyes of all vertebrates are stabilized relative to the optical environment. Exploratory or saccadic eye-turning could not occur except for this fundamental stabilization of the eyes to the ambient light. There is an invariant reciprocal relationship between compensatory eye-turning and head-turning. Hence, in the stimulus-array there will be an invariant reciprocal relationship between the projected form of the iris in the eye and that of the face for any fixed direction of gaze.

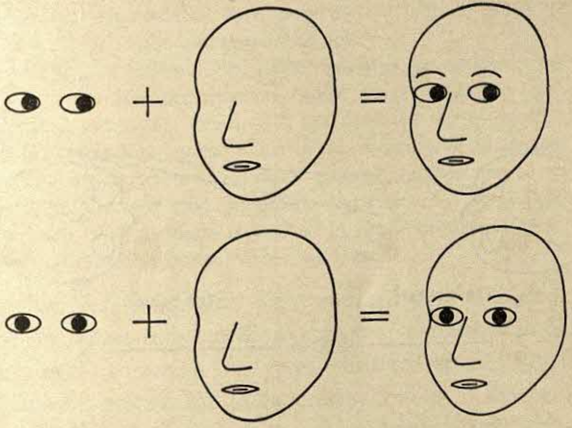


FIG. 2. THE SAME EYE-FORMS AS IN FIG. 1 SURROUNDED BY A FACE-FORM WHICH SHIFTS THE APPARENT DIRECTION OF THE EYES

Eyes formerly gazing to the right now seem to gaze forward, and those formerly gazing forward now gaze to the left.

From all this we are led to a more promising hypothesis about the stimulus for the perception of the direction of gaze: it is not the eye-form as such but the eye-form relative to the face-form. More exactly, the information is given by a constant relation of the asymmetry of the iris to the opposite asymmetry of the face. The eye-pattern can be variable in the manner of Fig. 1, but since this variation is reciprocal to the variation of the larger pattern around it, an invariant exists. The value of this invariant specifies the direction in which the eyes are looking.

In short, the problem is one of spatial perception. One perceives the orientation of the head in space just as one perceives the orientation of other mobile objects in the environment, and because one also registers the posture of the eyes in the head, one can perceive the *absolute* orientation of the eyes in space. So considered, the problem has implications for the classical puzzle of shape-constancy. According to our hypothesis, the stimulus for *being looked at* would be a nearly equal asymmetry of the two linked variables. Unequal asymmetry would specify a direction of gaze to one side or the other. It is this part of the hypothesis that we propose to test. Is such an invariant relation among forms detectable, and with what accuracy?

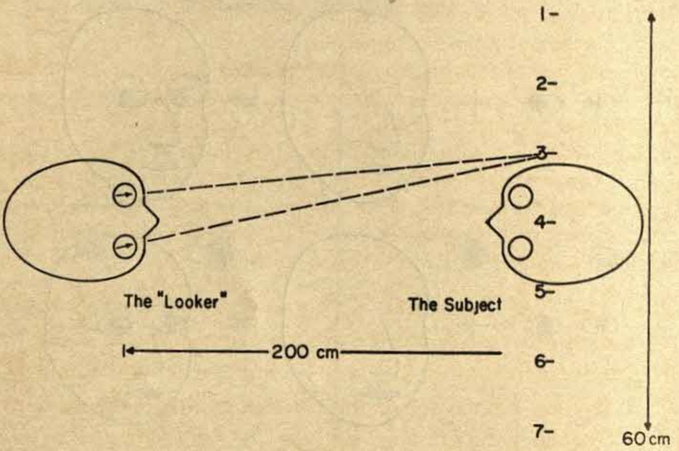


FIG. 3. VIEW OF THE EXPERIMENTAL ARRANGEMENT FROM ABOVE
The horizontal distance is shortened relative to the vertical.

THE EXPERIMENT

For a study of the psychophysics of the perception of the gaze of another, we might use as stimuli drawings of objects, photographs, models, or the postures of a real person. Since the experiment was exploratory, we chose the last. A student was employed and trained as an impersonal source of controlled social stimulation. At the request of *E*, she would adopt one of seven eye-postures in one of three head-postures relative to the *O*. The method of constant stimuli was used. The task of *O* was to judge whether he was or was not being looked at. How accurate would these judgments be? Specifically, we measured the smallest deviation of the looker's line of regard from the bridge of *O*'s nose that could just be noticed. We wished to be able to compare the acuity of this discrimination with other types of visual acuity.

The 'looker' had large brown eyes. She had been carefully drilled in the procedure to be used, and maintained the same passive facial expression on all trials. Her eyes had been tested and shown to be free of ocular imbalance, since a preliminary unpublished experiment by Wardwell had shown that even a slight phoria affected *O*'s judgments adversely. The looker sat opposite *O* at a distance of 2 m., in ordinary room-illumination—at a level less than optimal for other types of acuity.

The variable stimuli, the absolute orientations of her eyes in the room, were established by having her look steadily at one of seven fixation-points on or near the face of the *O*. A scale of seven numbered spots was marked on the wall just behind *O*'s head, as shown in Fig. 3. The middle spot was the bridge of *O*'s nose and the others were to the right or left of his head. The spots were 10 cm. apart, the angular separation being 2.9° for the central part of the scale at a distance of 200 cm. The seven eye-postures were presented at each of three head-postures: facing the *O*, turned 30° right, and turned 30° left.

The procedure was as follows: The looker assumed one of the 21 eye-fixations in accordance with a prearranged random order, maintaining the fixation steadily, until the judgment was made, usually 5-6 sec. *O* noted the appearance of her eyes and face and then made a judgment of *yes* or *no*, meaning "You are looking directly at me," or "You are not looking directly at me." *O* then turned away and the looker prepared for the next trial. Judgments were recorded by *E*. Each *O* made 25 judgments for each of 7 fixations in each of 3 head-positions, a total of 525. Six adult *O*s, all with presumably normal vision, were used.

Results. The 'yes' judgments were counted and tabulated. Some *O*s gave more than others did, but the accuracy of perception could be measured by the degree to which the 'yeses' clustered around Point 4, on the scale of 7 fixation-points, the point corresponding to the bridge of *O*'s nose. The frequency-distributions of 'yes' judgments over the scale were plotted for each of the three head-positions, for each *O* separately. Individual differences were not large; hence the *O*s were pooled and the three mean distributions are shown in Fig. 4. The *SD*s are 0.92, 0.80, and 0.93, reading from left to right.

It can be noted that for the frontal head-position the pooled distribution is centered very close to Point 4, indicating no constant error in the group as a whole. For the other head-postures the distributions are shifted significantly, showing consistent constant errors. The shift is in the direction of head-pointing. Evidently the perception of eye-pointing is somewhat influenced by the perception of head-pointing.

The *SD*s are a measure of the precision of judgments. They are all less than one step on the scale of fixation-points. The peaked distributions show a clustering of the 'yeses' around a certain stimulus-presentation which indicates acute perception of being looked at.

The *SD*s for the different head-positions are not significantly different from one another. This means that acuity for the perception of gaze is independent of head-pointing. We conclude that the absolute orientation of the eyes in space is detectable, although with some constant error. The invariant postulated as existing between the two reciprocal form-variables must be effective for perception or, if not this precise invariant, something equivalent to it.

Comparison with other types of acuity. We take the mean *SD* of the distribution of all 'yes' judgments in this experiment to be a measure of a threshold, the just noticeable deviation of the looker's line of regard from *O*'s nose. This deviation corresponds to a distance of about 9 cm. seen from 200 cm. It is roughly the distance from the nose to the edge of the face. The angular deviation of the looker's eyes is 2.8° for the *SD* with head front. The corresponding linear displacement of the looker's iris (the frontal displacement of the eyeball corresponding to a rotation of 2.8°) is less than 1 mm. This degree of asymmetry is slight. A displacement of 1 mm. at a distance of 2000 mm. is a very small angle of intercept, about 1 min. of arc. The minimal separation of the elements of a printed letter necessary for its identification is usually taken to be 1 min. of arc in the Snellen test. If it is remembered that the level of illumination on the looker's face was not as high as is customary on an acuity-chart, and that acuity depends on illumination, we may safely conclude that this type of acuity is at least as good as Snellen acuity.

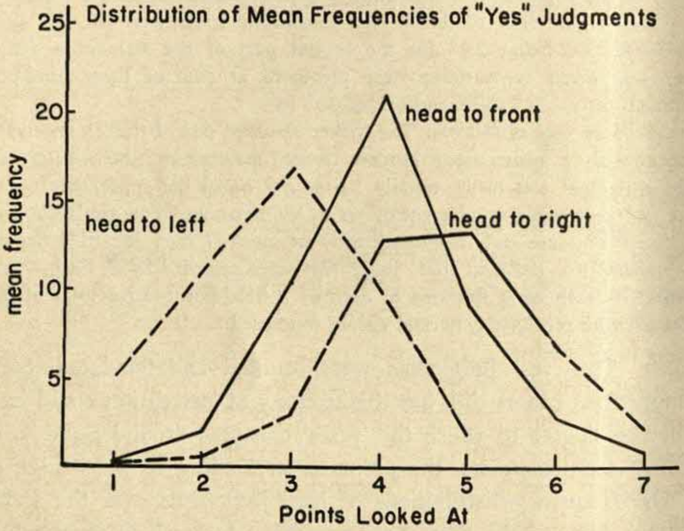


FIG. 4. POOLED DISTRIBUTIONS OF THE JUDGMENTS OF 'YES' AT SEVEN DIFFERENT EYE-POSITIONS FOR EACH OF THREE HEAD-POSTURES

The pattern-variable by which one detects the line of regard of a person is geometrically complex, as Fig. 2 shows. It is not a 'form' but a high-order relation between forms. Evidently this does not mean, however, that acuity for such optical information is any less than it is for some of the geometrically 'simpler' variables measured in standard acuity-tests. Very subtle properties of visual pattern are known to be stimuli for perception. When, as in our experiment, they specify the state of another person, they may come to be discriminated with the greatest efficiency possible for the visual receptive system.

Discussion and conclusions. This experiment should no doubt be repeated without using a live person as the source of stimulation, perhaps with a realistic model having movable calibrated eyes, or with a set of photographs, or with schematic drawings. Conceivably, our looker unconsciously betrayed the act of fixating the center of *O*'s face by some slight change in facial expression, despite training to the contrary. No such changes could be observed by *E*. Furthermore, if such a minimal cue had been utilized exclusively by the *O*s, it would have been as effective in the deviated head-positions as in the frontal head-position, and the constant errors would not have occurred. We believe that a real person can be used successfully to produce the stimuli for a psychophysical experiment. Nevertheless, a model or pictures would yield direct control by *E* of the stimulus-variable and would guarantee the elimination of all other cues.

The possibility of separating the ocular expression of the direction of attention from the facial expression of emotion at least has been demonstrated. This should be considered in future studies of the perception of human heads. The stimulus-variable corresponding to the direction of gaze, in fact, can be analyzed and experimentally controlled more easily than the stimulus-variables corresponding to emotional expressions. The families of elastic deformations of the face that we see as frowning, smiling, and the like are hard to dimensionalize, although Schlosberg has started to do so.[4] The rotations of the eyes and head, on the contrary, are mechanical movements to a close approximation, which are geometrically simpler. The elastic movements of the facial surface provide optical transformations of a group belonging to topology but not to perspective geometry.[5] The rotary movements of the eyes and head, being rigid instead of elastic, provide transformations corresponding to different perspectives of the same object. The stimulus-information is of a mathematically different order, and the corresponding perception is typically one of orientation, not deformation.[6]

---

[4] R. S. Woodworth, and Harold Schlosberg, *Experimental Psychology*, 1954, Ch. 5.

[5] Kai von Fieandt and J. J. Gibson, The sensitivity of the eye to two kinds of continuous transformation of a shadow-pattern, *J. exp. Psychol.*, 57, 1959, 344-347.

The results suggest that we have good discrimination for the line of gaze of another person, at least with respect to whether or not we are being looked at. The ability to read the eyes seems to be as good as the ability to read fine print on an acuity-chart, according to our first determination.

The variable error of judgment does not appear to be signficantly greater when the eyes look 'askance' at me than when they look directly at me, although a constant error is produced under the former conditions. This result is consistent with the hypothesis of a relational stimulus-variable for the perception of gaze.

Finally, we have shown the possibility of using a psychophysical method for studying a seemingly complex stimulus-variable—but a variable which conveys information for meaningful perception. All the various styles of looking that were listed at the beginning are properly meaningful in social life. We have isolated only one of the variables in this complex, the temporary line of gaze. Other qualities of looking behavior might be amenable to experimental control if the attempt were made. The study of the expressive reactions of men and animals in general, the basis of social interaction, might profit by this analytical approach.

---

[6] J. J. Gibson, and E. J. Gibson, Continuous perspective transformations and the perception of rigid motion, *J. exp. Psychol.*, 54, 1957, 129-138.

THE RELATION BETWEEN CATEGORY-RATING AND MAGNITUDE-ESTIMATION JUDGMENTS OF OCCUPATIONAL PRESTIGE

By SIDNEY I. PERLOE, Haverford College

The studies to be reported here were designed to determine whether the curvilinear relationship found between sensory scales based on judgments of differences and sensory scales based on judgments of ratios also would be found with scales of meaningful stimuli.[1] This relationship has been attributed to differences in *S*'s ability to discriminate one stimulus-magnitude from another. At the low end of prothetic dimensions, discrimination is said to be fine, with each judgmental category of the difference-scale covering only a small segment of the continuum; at the high end, it is poor, resulting in broader categories. If ratio-judgments are unaffected by the region of the continuum from which the stimuli are drawn, the difference-scale units, will encompass more ratio-scale units at the lower than at the upper ends of prothetic continua. It also has been proposed that the relationship between ratio-scale and difference-scale values approaches linearity when the latter are plotted against the logs of the former.[2] Stevens and Galanter believe that the attainment of linearity depends upon the spacing of the stimuli being judged.[3]

EXPERIMENT I

Although all the specific comparisons referred to above involved dimensions with clearly specified physical correlates, there is no reason to believe that all prothetic continua must have known physical correlates.[4] This experiment was designed simply to examine the relationship between magnitude-estimations (ratio) and category-ratings (difference) of the prestige commanded by various occupations. Occupational-prestige appears to meet the conceptual requirement of a prothetic continuum, *i.e.* it seems to vary to a great extent in terms of more or less rather than in terms of quality.[5] One might, therefore, expect that a comparison of scales of occupational prestige arrived at through ratio- and difference-methods would yield the same curvilinear relation that has been found in previous studies of prothetic sensory continua.

---

* Received for publication March 22, 1962. This work was supported in part by NSF Grants G-13023 and G-19253. The author is indebted to Dr. Robert Abelson of Yale University for helpful criticism, and to Doris Donovan, Simon Rich, Howard Stick, and Thomas Warne for their help in collecting and tabulating the data.

[1] Eugene Galanter and Samuel Messick, The relation between category and magnitude scales of loudness, *Psychol. Rev.*, 68, 1961, 363-373; S. S. Stevens and Eugene Galanter, Ratio scales and category scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1957, 377-411; W. S. Torgerson, Quantitative judgment scales, in Harold Gullikson and Samuel Messick (eds.), *Psychological Scaling: Theory and Applications*, 1960, 21-31.

[2] Galanter and Messick, *op. cit.*, 363-373; C. E. Helm, Samuel Messick and L. R. Tucker, Psychological models for relating discrimination and magnitude estimation scales, *Psychol. Rev.*, 68, 1961, 167-177; Torgerson, *op. cit.*, 21-31.

[3] Stevens and Galanter, *op. cit.*, 377-411.

Method. Four groups of 10 *S*s each were studied in a 2 × 2 design. There were 36 men and 4 women, all undergraduates. Each *S* judged a list of 100 occupations twice. One set of judgments was made on a 7-category rating scale and one set was made by magnitude-estimation. The occupations judged with a given technique were listed on two pages, preceded by a page of instructions. The instructions for magnitude-estimation asked *S* to assign a number to each occupation which was proportional to the amount of prestige possessed by the occupation. This instruction was elaborated to make clear the meaning of proportional assignment of numbers. *S* was allowed to start with any number he wished, *i.e.* no modulus was presented.

The occupations were drawn from a larger list, previously used by McGarvey.[6] The selection was made on the basis of judgments by pretested *S*s. It was aimed at obtaining occupations drawn from all parts of the available range and for which there was a relatively small amount of interjudge variability. As in McGarvey's study, *S*s were instructed to judge in terms of the prestige accorded to an occupation by a majority of the population. Two different random stimulus-orders were used, each *S* making one set of judgments with each order. One pair of groups made magnitude-judgments first and rating-judgments second, while the other pair made their judgments in the reverse sequence. One group in each pair received the two random lists in one sequence, while the other received them in the reverse sequence.

Results. In response to a question asked after the completion of both sets of judgments, it was discovered that many *S*s had spontaneously placed an *a priori* upper limit on the range of values they used in their magnitude-estimations. The responses of *S*s who limited the range of their magnitude-estimations (limited-range *S*s) and the responses of *S*s who did not (open-range *S*s) were analyzed separately. There were 16 open-range *S*s and 24 closed-range *S*s.

---

[4] Stevens, On the psychophysical law, *Psychol. Rev.,* 64, 1957, 153-181; Measurement, psychophysics, and utility, in C. W. Churchman and Philburn Ratoosh (eds.), *Measurement: Definitions and Theories,* 1959, 18-63.

[5] Stevens mentions the possibility that some dimensions may include both prothetic and metathetic (qualitative) variation. In the present case all that is required is that occupational prestige allow enough prothetic variation to generate the expected relationship.

[6] H. R. McGarvey, Anchoring effects in the absolute judgment of verbal materials, *Arch. Psychol.,* 39, 1943, (No. 281), 1-86.

The relationship between the two sets of scale-values was examined by plotting the mean rating given by a group of *Ss* to each of the 100 occupations against the corresponding median magnitude-estimation. The latter were spaced logarithmically to check the view that the values of the difference-scale are approximately linearly related to the logs of ratio-scale.

The responses of the open-range *Ss*, shown in Fig. 1 A, revealed the expected relationship between the two types of scale.[7] The fact that the limited-range *Ss* in the first study did not respond in the expected manner need not be interpreted as negative evidence, because these *Ss* did not fulfill the requirements of the task of magnitude-estimation.[8] By placing a limit on the range of values they allowed themselves to use, they essentially converted the task of magnitude-estimation into one of category-rating, with an expanded number of units. The curvilinear relation found between the magnitude-estimations of open- and limited-range *Ss*, presented in Fig. 2, supports this view of the limited-range responses.

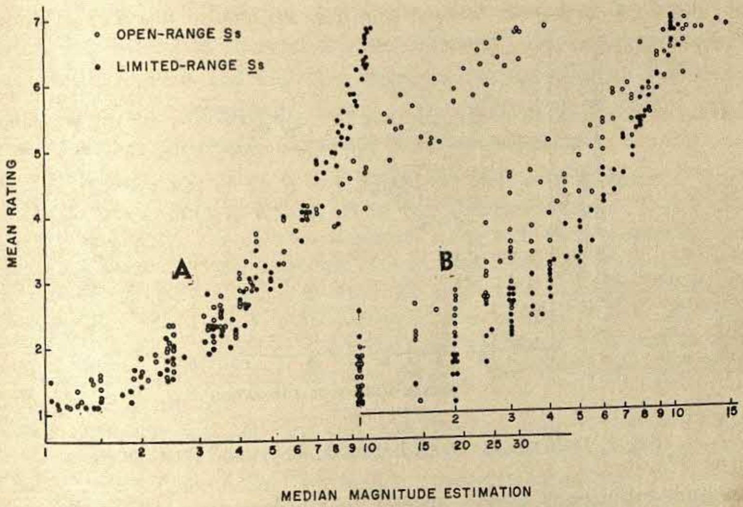


FIG. 1. MEAN CATEGORY-RATINGS VS. MEDIAN MAGNITUDE-ESTIMATIONS FOR OPEN- AND LIMITED-RANGE *Ss*

Plot A is based on a within-*Ss* comparison, using all *Ss*.
Plot B is based on a between-*Ss* comparison.

An inspection of the data revealed that the expected relationship occurred more strongly among open-range *Ss* making the category-ratings first and the magnitude-estimations second than among open-range *S*s performing the tasks in the reverse order. To eliminate the effects of task-order, the magnitude-estimations of the *S*s who performed them first were compared with the ratings of the *S*s who performed the category-ratings first. The former group had 4 open- and 16 closed-range *S*s. As seen in Fig. 1 B, the shape of the relationship between the *S*s of the groups is essentially the same as in the comparison within the *S*s. The major difference between the two plots is that open-range *S*s, who made their magnitude-estimations second, spread their judgments over a wider range of values than those who made them first.

---

[7] A plot of the mean ratings against the logarithmically spaced mean magnitude-estimations was essentially the same as the plot presented in Fig. 1A.

[8] Stevens, *op. cit.*, *Psychol. Rev.*, 1957, 64, 153-181; Churchman and Ratoosh, *op. cit.*, 18-63.

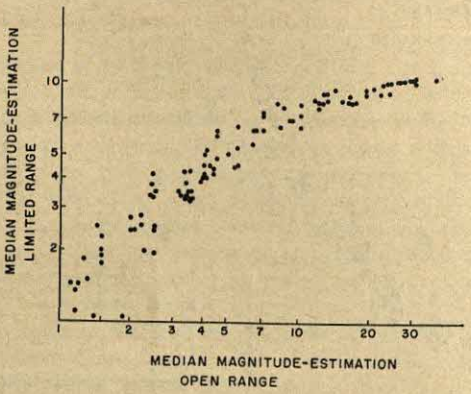


FIG. 2. LIMITED- VS. OPEN-RANGE MAGNITUDE-ESTIMATIONS

EXPERIMENT II

The second study was designed to test the following hypotheses based on proposals of Stevens and Galanter.[9] They are: (a) category-ratings will be related curvilinearly to corresponding magnitude-estimations regardless of the spacing of the stimuli along the judgmental dimension; (b) the relationship will be less curvilinear when stimulus-arrays of predominantly high magnitude are judged than when arrays of predominantly low magnitude are judged. The rationale for the second hypothesis is that *S*s expect stimuli from all parts of the stimulus-range to be presented equally often and therefore attempt to use each of the response-categories of the rating scale equally often. When the stimuli are unequally distributed along the judgmental dimension, *Ss* are believed to make their categories broader in the emptier regions and narrower in the more crowded regions in order to appoximate the ideal of placing the same number of stimuli in each category. Stevens and Galanter further imply that magnitude-estimations are not affected by stimulus-spacing, at least not as much as category-ratings are. Consequently, when the stimuli are concentrated at the low end of a prothetic continuum, the tendency toward category-equalization works to increase the already extant curvilinearity. When the stimuli are concentrated at the high end, the tendency works to decrease the curvilinearity. The first hypothesis is a test of the Stevens and Galanter assertion that the effect of differences in discriminability of stimuli drawn from different parts of the judgmental dimension outweighs the effects of stimulus-spacing.

---

[9] Stevens and Galanter, Ratio scales and category scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1967, 379-380.

Method. Eight groups of 11 *Ss* each judged the prestige of occupations. Except for two graduate students in classics, all *Ss* were undergraduate men. Each *S* made two sets of judgments, one on a seven-category rating scale and the other by direct magnitude-estimation. In the latter task, *S* was given a standard occupation in the middle of the occupational range to which the number 70 was assigned.[10] Explicit instructions also were given to avoid placing a lower or upper limit on the values they assigned. Four of the groups (high bias) received lists of 75 stimuli in which 50 were high in prestige and 25 were low; the other four (low bias) received lists with the reverse distribution.

Two groups with 'high-bias' and two with 'low-bias' made the category-ratings first and the magnitude-estimations second; the remaining four groups performed the tasks in the opposite order. Two random sequences were prepared for each set of occupations. The order in which the sequences were used was combined factorially with task-order in each of the bias-conditions.

The determination of the prestige of each occupation (required for the construction of the two biased sets) was made on the basis of previously collected judgments. A technique developed at the National Opinion Research Center was used to convert the judgments into prestige-scores.[11] The scores ranged from 20-100. In the present study, scores below 60 were considered low and scores of 60 and above were considered high. The range of prestige-values in the two sets was the same. The only difference was in the relative number of high- and low-prestige occupations. The two lists contained 43 occupations in common.

Results. The first hypothesis stated that the relationship between the category-ratings and magnitude-estimations would be curvilinear, regardless of the spacing of the stimuli, *i.e.* category-units at the low end of the prestige-dimension would encompass fewer magnitude-estimations than those at the high end.[12] An inspection of Fig. 3 A reveals that this was indeed the case.

---

[10] Due to an oversight, the standards for the high- and low-bias lists were not the same. The high-bias standard was *accountant*, while the low-bias standard was *auditor*. Both had prestige-scores of 62.5. (See below for discussion of prestige-scores.)

[11] National Opinion Research Center, Jobs and occupations :A popular evaluation, in R. Bendix and S. M. Lipset (eds.), *Class, Status and Power: A Reader in Social Stratification*, 1953, 411-426.

The second hypothesis was that the relationship would be more curvilinear when *Ss* judged low-bias lists than when they judged high-bias lists. The lack of difference between the high- and low-bias plots fails to support this hypothesis. The expected difference was based on the assumption that *Ss* would tend to assign the same number of occupations to each category. An inspection of the data showed wide individual differences in the degree to which this assumption was met. To examine these differences more precisely, the variance of the frequencies with which the several rating-scale categories were used was computed for each *S*. The entire sample was split at the median variance. *Ss* below the median were considered to manifest high equalization-tendencies, while *Ss* above the median were treated as having low equalization-tendencies.

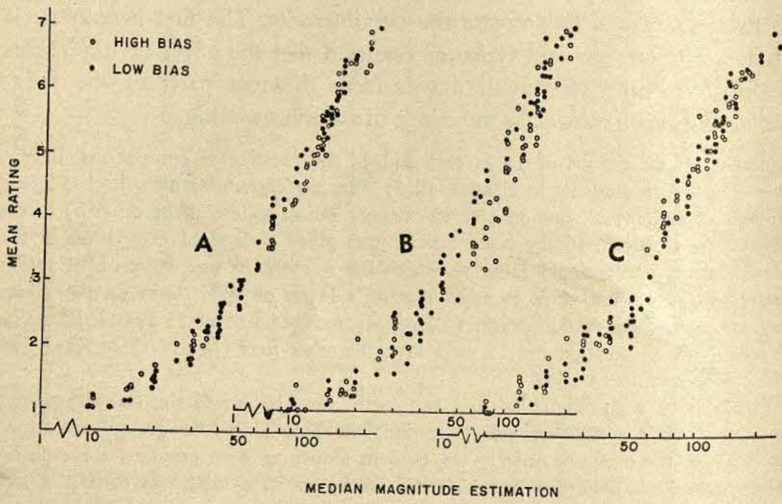


FIG. 3. MEAN CATEGORY-RATINGS VS. MEDIAN MAGNITUDE-ESTIMATIONS FOR HIGH- AND LOW-BIAS LISTS

Plot A is based on the judgments of the entire sample. Plot B is based on the judgments of the high-equalization *Ss*. Plot C is based on the judgments of the low-equalization *Ss*.

---

[12] Since the mean category-ratings were related to the mean and the median magnitude-estimations in the same way, only the relationships involving the median magnitude-estimations will be presented.

The second hypothesis then was tested separately for the high- and low-equalization *S*s. There were 25 high-equalization *S*s in the high-bias group and 19 in the low-bias group. It can be seen in Fig. 3 (B and C) that the data for the high-equalization *S*s are consistent with the second hypothesis, while the data for the low-equalization *S*s are not.

Although the failure to confirm the hypothesis might have been due to lack of a sufficient difference between the two distributions, the observed context-effect (see Fig. 4 and related discussion) and the occurrence of the expected result with the high-equalization *S*s argue against this view. The lack of support for the hypothesis seems rather to be due to the inadequacy of the two assumptions on which it was based, the tendency to use categories equally, and the immunity of magnitude-estimations to the effects of stimulus-spacing. It should be noted, however, that at least partial support is provided for the latter assumption by the fact that the magnitude-estimations of the high-equalization *S*s were less affected by differences in stimulus-spacing than were their ratings.

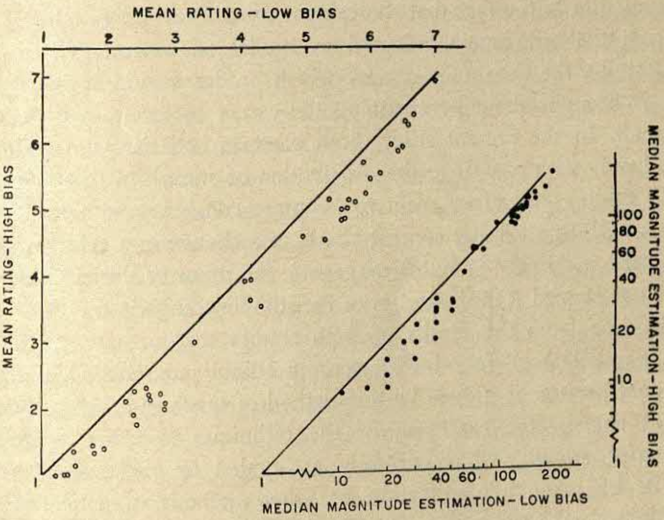


FIG. 4. COMPARISONS OF JUDGMENTS OF OCCUPATIONS COMMON TO THE HIGH- AND LOW-BIAS LISTS

Plot A compares the category-ratings made in the high-bias context with those made in the low-bias context. Plot B makes a similar comparison for the magnitude-estimations.

The data raise two major questions about the effects of stimulus-spacing on the relationship under study. First, what conditions are responsible for the variation in the degree of equalization which occurred? Secondly, how widespread are these conditions; that is, to what extent will they be found over the range of situations in which psychologists attempt to scale stimuli?[13]

The presence of many occupations common to lists of both high- and low-bias permitted a comparison of the judgments made to identical stimuli in the different contexts (see Fig. 4). In both cases, higher values were assigned to occupations when they occurred in contexts of low-bias than when they occurred in contexts of high-bias. The major import of this result lies in the fact that the present procedure was somewhat different from the one usually adopted in studies of context-effects. In the typical study, the change in context which produces judgmental displacements is brought about by changing the extent or location of the range of stimuli. In the present study, both contexts had the same range and differed only with respect to the distribution of stimuli in that range. The present findings therefore seem to be more consistent with a theory of judgment which attributes context-effects to a change in a reference-standard located within the stimulus range, *e.g.* the mean or the median, rather than to one located at the boundary of the stimulus-range.[14]

It is clear from the results of both studies that regular relationships which parallel those found with sensory stimuli are obtained when occupational prestige is judged by the method of magnitude-estimation. The findings suggest the extension of the technique to other domains of meaningful stimuli, which currently are scaled by difference methods. The use of magnitude-estimation or other methods of ratio-scaling in the judgment of socially relevant material is unfortunately complicated by the fact that many of these vary along bipolar dimensions. While variation within a given region can be measured conveniently with ratio-methods such as magnitude-estimation, problems arise when material from the two regions of the continuum are judged together. How, for example, should *S* assign a number proportional to the liberalism of a senator he believed to be extremely conservative? Should he assign a zero? Should he assign a negative value which is proportional to the senator's distance from the boundary between liberalism and conservatism? Should he first sort the senators into liberal and conservative categories and then perform the magnitude-estimation task separately for degree of liberalism and degree of conservatism? It is hoped that future development of the ratio-methods will produce solutions to these problems and facilitate their use in the scaling of meaningful materials.

---

[13] Allen Parducci, R. C. Calfee, L. M. Marshall, and L. P. Davidson, Context effects in judgment: Adaptation level as a function of the mean, midpoint, and median of the stimuli, *J. exp. Psychol.*, 60, 1960, 65-77.

[14] *Ibid.,* 65-77.

SUMMARY

The studies reported here were designed to determine whether scales produced by magnitude-estimations and category-ratings would have the same relationship to each other in judgments of occupational prestige as they do in judgments along prothetic sensory continua. The first study tested the hypothesis that the rating-scale categories at the high-prestige end of the dimension would encompass more magnitude-estimation units than those at the low end. The expected, approximately logarithmic, relationship was obtained with *S*s who did not place an *a priori* upper limit on their magnitude-estimations.

The second study was designed to test the hypotheses (a) that category-ratings of occupational prestige would be curvilinearly related to the corresponding magnitude-estimations, regardless of the stimulus-spacing; and (b) that the degree of curvilinearity would be less when arrays of predominantly high-prestige occupations were judged than when predominantly low-prestige arrays were judged. The data supported the first hypothesis, but not the second, although the judgments of the *S*s who used the rating-categories equally often were consistent with the second hypothesis. It also was found that both types of judgment of a given list were displaced away from the crowded end of the dimension.

In general, the results of the two studies suggested that ratio-scaling methods, particularly magnitude-estimation, might profitably be extended to the judgments of meaningful materials.

THE CONSTANCY OF STEREOSCOPIC DEPTH

By HANS WALLACH, Swarthmore College,
and CARL ZUCKERMAN, Brooklyn College

It is one of the elementary facts of stereoscopic vision that retinal disparity represents depth quantitatively. The greater the distance in depth between two points on an object, the greater will be the disparity with which their images are given in the two eyes, and, up to a point, the greater will be the perceived depth-interval between them. Nevertheless, just as the size of a retinal image does not depend alone on the size of the corresponding object, the amount of retinal disparity between two points does not depend exclusively on their distance in depth; in both cases, the distance of the object from the eyes is important. There should therefore be a constancy-problem in stereoscopically perceived depth which corresponds to the problem of size-constancy: If the perceived-depth between two points on an object is to correspond to the distance in depth between the two points on the physical object, the distance of the physical object from the eyes must, in some fashion, be taken into account.

The problem of the relation between perceived depth and the distance of the depth-interval from the eyes has attracted surprisingly little attention in the last half century, considering the great interest that exists in stereoscopic vision. Older experiments, to which references may be found in Linschoten,[1] did demonstrate a considerable degree of stereoscopic depth-constancy. In this paper, we explore some of its properties and attempt to show how they serve to explain some familiar phenomena.

Although the problem of stereoscopic depth-constancy and the problem of size-constancy owe their existence to the same variable, namely, the distance of the object from the eyes, they differ in quantitative terms. Whereas the size of the retinal image decreases directly with the distance of the corresponding object from the eyes, the amount of retinal disparity is inversely proportional to the square of the distance.[2] It follows that

---

* Received for prior publication March 2, 1963. The work presented here was supported by Grant G-9599 from the National Science Foundation.

[1] J. Linschoten, *Structuranalyse der binocularen Tiefenwahrnehung,* 1956, 180.

[2] C. H. Graham, Visual perception, in S. S. Stevens (ed). *Handbook of Experimental Psychology,* 1951, 888; K. N. Ogle, *Researches in Binocular Vision,* 1950, 136.

different mechanisms must underlie the two constancies, if the perception of stereoscopic depth operates more or less adequately.

That disparity should be inversely proportional to the square of the distance of the depth-interval from the eyes can easily be made clear. There are two reasons why disparity should vary with distance: (1) Inasmuch as disparity consists in small differences in the width of the retinal images in the two eyes, it must decrease in proportion to the distance of the object from the eyes as do retinal images themselves. (2) Disparity comes about because the two eyes view an object from slightly different directions. The greater this difference in viewing direction, the greater will be the disparity that results from a given depth-interval (Fig. 1). The farther the object is from the eyes, the less the two directions differ from each other, and the smaller should be the disparity that a depth-interval of a given amount will cause; in fact, disparity should

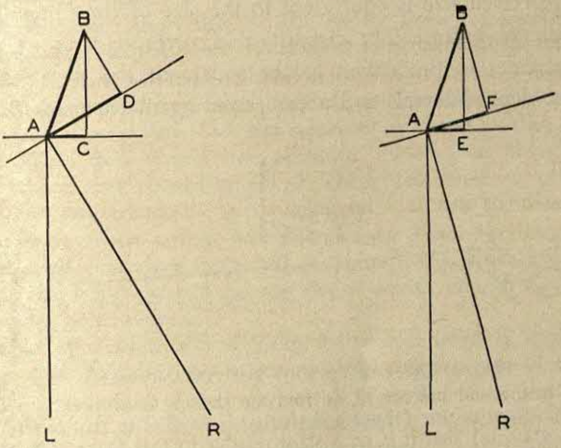


FIG. 1. RETINAL DISPARITY AS A FUNCTION OF DISTANCE

A and *B* are two points in the horizontal plane which are arranged in depth in relation to *O*. *AL* is the direction from *A* to the left eye and *AR* the direction to the right eye. On the left, *AC* is the projection of *AB* on the frontal plane through *A* of the left eye, *AD* its projection on the frontal plane of the right eye. The difference in length of these two projections is the immediate cause of the disparity when *AB* is given. On the right, the distance of *AB* from the eyes is assumed to be doubled, which diminishes the difference in the directions from which the two eyes view *AB* to one half (compare the angles *LAR*). As a consequence, the projections of *AB* on the frontal planes of the two eyes, *AE* and *AF*, are more nearly equal; they differ in length only half as much as do *AC* and *AD*. This change in the relative width of projections of depth-intervals with changing viewing direction is one of the causes for the dependence of disparity on distance.

decrease, for this reason, in proportion to the distance of the object from the eyes. Since these two factors are independent of each other, disparity should decline with the square of the distance.

Let us briefly compare stereoscopic depth-constancy with size-constancy. The size of an object's retinal image decreases in proportion to its distance from the eyes, and the perceptual process, compensating for this size-loss in the stimulus-situation, operates more or less according to Emmert's law: The perceived size is equivalent to the size of the image multiplied by the registered distance. The disparity that corresponds to a given depth-interval decreases in proportion to the square of the distance from the eyes of that depth-interval, and a compensating mechanism that leads to complete constancy would therefore have the perceived depth be dependent on the given disparity multiplied by the square of the distance.

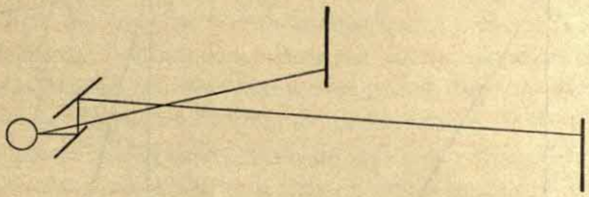


FIG. 2. LATERAL VIEW OF THE ARRANGEMENT

Only the right eye and one set of mirrors are shown. Convergence is changed by so turning the upper mirror (about a shaft that is parallel to that mirror and to the plane of the drawing) that it is on a slant relative to that plane. The vertical rods are shown as they are used for adjusting convergence to half the object's distance.

Probably the clearest demonstration of the compensating process underlying stereoscopic depth-constancy can be obtained through an arrangement that permits manipulation of the cues for distance of an object from the eyes while keeping its actual distance, and therefore the disparity, constant; this should, of course, result in changes of perceived depth. In such an experiment, three-dimensional wire forms were observed either directly or through lenses and a mirror-arrangement which changed accommodation and convergence. The effect of the artificial alteration of these distance-cues on the apparent depth of the wire form could be gauged by comparing depth-estimates made under the two conditions of observation.

A sketch of the mirror-arrangement used to manipulate convergence is shown in Fig. 2. The ocular mirrors are so fixed in a 45° slant in *O*'s sagittal plane that they reflect vertical beams into *O*'s eyes. The objective mirrors are suspended by two joints. One joint lets the mirror turn about a horizontal frontal-parallel shaft and makes tilting it in *O*'s vertical sagittal plane possible; it is used to make the viewed object appear at its true elevation above the table. The shaft of the other joint is parallel to the mirror-surface and lies in the vertical sagittal plane; these latter joints are used to make the object appear straight in front and to manipulate the convergence with which it is viewed. To adjust this convergence, a thin vertical rod is set on the table in the place where the experimental object will eventually be; this rod is seen only through the mirror-arrangement. Another vertical rod is so suspended from above that it is visible through the gap between the ocular and the objective mirror; it is placed at the distance of the intended convergence. By so adjusting the objective mirrors that for each eye the lower rod appears as an extension of the upper rod, the two rods are made to be viewed with the same convergence and will then appear to be at the same distance.

Two wire-forms, both shaped like four-sided pyramids with nearly square bases but differing in size, were used. Each was supported by a shaft attached to one edge of the base. When these objects were presented, the base was always in frontal orientation and the apex pointed toward *O*. The critical object was placed on the table and was viewed through the mirror-arrangement. Its base had an edge of 12.5 cm., and the distance from apex to base was 12.4 cm. The other wire-pyramid was half as large in all dimensions and served as control object. It was suspended from its shaft above the larger pyramid and was viewed directly through the gap between the ocular and the objective mirror.

The *O*s gave depth-estimates by adjusting the variable length of a thin brass rod they held in their hands to the apparent distance between apex and base of a pyramid. This procedure of obtaining judgments of depth had previously been used successfully.[3] Perceived size also can be expected to be altered when cues for distance are changed, and size-estimates, using the length of a diagonal of the base, were obtained along with the depth-estimates. This made it possible to compare stereoscopic depth-constancy with size-constancy under identical conditions.

The critical wire pyramid was placed 133 cm. from *O*'s eyes, with the detour by way of the objective mirrors taken into account. The mirrors were so adjusted that *O* viewed the wire-form with a convergence for half the real distance (66.5 cm.), and he wore —0.75-diopters lenses, enabling him to see the wire form with the eyes accommodated for the same half-distance.[4] Under the assumption of complete constancy in accordance with the altered distance-cues, *O* should perceive the large pyramid under these critical conditions with half the normal size and with 25% of the normal depth. Since retinal image-size and disparity remained unchanged, and convergence and accommodation represented half the real distance, perceived size should be expected to become reduced by a factor of two and perceived depth by a factor of two squared. These expectations could be tested by comparing the size- and depth-estimates obtained for the critical pyramid with estimates obtained for the control pyramid which was also placed 133 cm. from *O*'s eyes and was viewed directly and without the lenses, *i.e.* with veridical convergence and accommodation. Since the control pyramid was half as large as the critical one, complete constancy under both conditions of viewing would cause the bases of the two pyramids to appear equal in size, and apparent depth should be half as great for the critical as for the control pyramid (Table I, column IV).

---

[3] Hans Wallach, M. E. Moore, and Linda Davidson, Modification of stereoscopic depth perception, this JOURNAL, 76, 1963, 191-204.

[4] When a real object is moved from a distance of 133 cm. to a distance of 66.5 cm., accommodation should increase by 0.75 diopters. The same change should take place when an object that stays at a distance of 133 cm. is viewed through a —0.75 diopter lens for which accommodation has to compensate.

The 18 *O*s who were studied here first viewed the larger pyramid and gave an estimate of depth and one of size, while the smaller pyramid was hidden by a screen, then the larger pyramid was covered and they made corresponding estimates for the smaller pyramid.

Results. The first and third columns of Table I give the results. They show that manipulation of convergence and accommodation alone produce a fair degree of size-constancy as well as of stereoscopic depth-constancy. Note in this connection the small confidence-limits, which nowhere exceed 15%. A comparison of the ratios formed by the critical and the control-estimates with the expected ratios shows that the deviation from ideal constancy was about the same for size and for depth, and amounted to about 25%. No generalization should be made, however, from this approximately equal lag in constancy. The deviation from ideal size-constancy is ordinarily much smaller. Ours is undoubtedly due to the fact that convergence and accommodation were the only distance-cues available. The parallel between the two constancies may well cease when further distance-cues are introduced. The results are, nevertheless, impressive when the effectiveness of convergence and accommodation is considered. They contradict the widely held view that accommodation and convergence do not serve as reliable cues for distance.[5]

TABLE I

RESULTS OF THE EXPERIMENT WITH WIRE-FORMS

	Mean estimates and confidence-limits at the 1% level		Ratio (critical/control)		
	Critical condition	Control condition	Experi-mental	Complete constancy	No constancy
Size	8.88±0.82	7.13±0.59	1.245	1.0	2.0
Depth	4.07±0.40	6.43±0.82	0.633	0.5	2.0

(1) Behavior of the anaglyph. A simple and interesting method of demonstrating stereoscopic depth-constancy involves the use of anaglyphs, in which the left-eye and the right-eye view of an object or a scene are printed in different colors, one superimposed on the other. Viewed through colored spectacles so chosen that the right eye sees only the right-

---

[5] See E. G. Heinemann, Endel Tulving, and Jacob Nachmias, The effect of oculomotor adjustments on apparent size, this JOURNAL, 72, 1959, 32-45.

eye view and the left eye only the left-eye view, the anaglyph produces the stereoscopic effect. An anaglyph has useful properties: It is, as it were, an outward projection of the given retinal disparities, for the horizontal distances between corresponding left-eye and right-eye contours represent those disparities.

We found that when the distance of an anaglyph from the eyes is changed, stereoscopic depth is altered—increasing this distance will enhance perceived depth—and this is a manifestation of stereoscopic depth-constancy. To understand this we have to recall the two reasons for the decline of disparity with the square of an object's distance from the eyes. The first one, that disparity, which consists in differences in width of the retinal images of an object in the two eyes, must decrease with the distance of the object as do the images themselves, applies to the anaglyph also. As its distance from the eyes increases, its retinal image decreases in size and so also do the projections of the distances between left-eye and right-eye contours. The second reason for a decline of disparity with distance under normal conditions has to do with the way surfaces or pairs of points that occupy depth project on the frontal plane of each eye when distance is varied (Fig. 1). For a given depth-interval, the difference in the width of the projection, which, in turn, causes the disparity, varies inversely with distance. In the anaglyph, it is these projections that are given rather than the depth-intervals as such, and there is no reason for these projections to vary when the distance between the anaglyph and the eyes is changed. With only one of the two reasons for the change of disparity with distance operating, the disparity caused by an anaglyph varies in proportion to the distance only.

We have already stated that, inasmuch as disparity is inversely proportional to the square of a depth-interval's distance from the eyes, a compensating perceptual process must enhance perceived depth in proportion to the square of the distance, if that depth is to remain completely constant. Since, in the case of the anaglyph, disparity decreases in proportion to the distance only, a perceptual process that enhances perceived depth in proportion to the square of the distance does more than make up for the 'distance loss': While disparity declines in proportion to the distance only, the perceptual process enhances depth in proportion to the square of the distance. The net result is that perceived depth in the anaglyph should increase ideally in proportion to an increase in distance.

A measurement of the actual effect of distance-change on anaglyph-depth was taken. The *Os* were asked to reproduce the apparent depth-interval between two particular contours in an anaglyph with a caliper.

Each one of 39 *Os* reproduced the same depth-interval twice, once when the anaglyph was 18 in. from the eyes and again when that distance was doubled. For the 18-in. distance, the mean of the caliper-settings was 2.07 cm. and for the 36-inch distance it was 3.45 cm. The ratio of 0.6 between these means is to be compared with the ratio of 0.5 which would denote complete constancy. Though small, this deviation from ideal constancy turns out to be highly reliable ($p < 0.001$).

The dependence of the depth perceived in an anaglyph on the distance of the anaglyph from the eyes was also observable in three-dimensional motion pictures, for the stereoscopic effect here was produced by an anaglyph technique. Variations in apparent depth with a change in *O*'s distance from the screen were here quite startling, because of the great range over which distance could be varied.

(2) Perspective-cues in stereoscopic depth-constancy. With the help of the anaglyph technique, we were able to show that not only oculomotor cues for distance operate in stereoscopic depth-perception, but perspective cues also, as is the case in size-perception. Proving this is difficult because with stereoscopic depth the main issue, binocular regard as such cannot be avoided. The demonstration made use of a pseudoscope, an instrument that reverses stereoscopic depth, to create a discrepancy between binocular and perspective distance-cues. Four identical anaglyphs were prepared. Two of them were suspended in the upper part of the pseudoscope's field, one of these at a 3- and the other at a 5-ft. distance from *O*'s eyes. The other two were placed on the table below which was covered with a plaid tablecloth to strengthen perspective-cues. They also were 3 and 5 ft. from *O*'s eyes. Viewed through the pseudoscope, all four anaglyphs showed the expected reversal of depth. In the case of the suspended anaglyphs, the distances at which they were perceived were also reversed; the nearer one appeared to be more distant than the objectively farther one. Now we come to the first essential point: The nearer anaglyph which, due to the pseudoscopic distance-cues, appeared to be farther away, showed a greater depth. Although the same pseudoscopic distance-cues were present for the analyphs that rested on the tablecloth, their apparent distances from *O* were not reversed. As is normal for this particular contradictory arrangement of cues, the perspective-cues for distance prevailed over the binocular ones, and the objectively more distant anaglyph appeared to be farther away. There was again a difference in depth displayed by the two anaglyphs, but here the objectively farther anaglyph showed the greater depth. Perceived depth varied in accordance with the prevailing distance-cues, the

perspective ones, which thus proved to operate in stereoscopic depth-constancy.

(3) Effect of optical magnification. The dependence of stereoscopic depth on distance also largely explains the familiar fact that a scene viewed through binoculars appears to have reduced depth.[6] Binoculars, like other telescopes, magnify the retinal image of a viewed object. For various reasons, this magnification results in a shortening of the perceived distance rather than enhanced perceptual size. Familiar size undoubtedly plays a role here, particularly when one looks directly at the scene before viewing it through the glasses; but convergence also probably is responsible for the diminished apparent distance, for, with binoculars, magnification increases the angle of convergence with which an object is viewed, as well as its retinal image (Fig. 3).

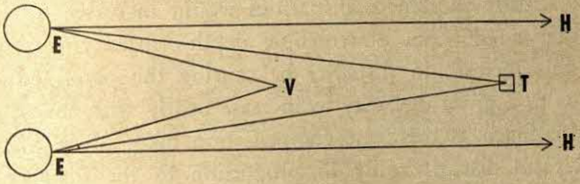


FIG. 3. EFFECT OF MAGNIFICATION ON CONVERGENCE

EH are the lines of sight when the eyes are converged for infinity. The angles *TEH* are the convergence-angles for an object *T*. Binoculars will magnify the angles *TEH* in proportion to their power, producing new convergence-angles *VEH*, which serve as cues for a diminished distance of the object. In every convergence-angle, one leg must coincide with *EH*.

---

[6] Among previous interpretations of reduced depth due to magnification, J. M. Vanderplas (On the flattening-effect of optical magnification, this JOURNAL, 73, 1960, 473-478) proposed to explain the effect in purely optical terms. He claimed that an object's final image formed by a telescope has a location "in optical space" which, in relation to the size of that image, might account for the effect. However that may be, Vanderplas failed to explain how the location of an image "in optical space" can be given to the eye.

S. Howard Bartley pointed out (*Beginning Experimental Psychology*, 1950, Chap. 15), that the retinal images of a near and a far surface of a solid object do not only increase in size when the object is brought nearer to the eye, but also change the *ratio* of their sizes. Optical magnification also increases image-sizes but does not change that ratio. If the size-ratio of the surfaces themselves were given, the image-size ratio produced by magnification could indeed lead to the flattening effect, but Bartley does not explain how that objective ratio is given. It is conceivable that, in the case of an orthogonal object and with an implicit assumption of orthogonality operating in perception, the ratio of image-sizes might, in fact, be the basis for the perception of depth, and therefore cause the flattening effect. If this were the explanation, the flattening effect should be restricted to orthogonal objects, but that is not the case. We were able to demonstrate with depth-estimates a strong reduction in depth when the wire pyramid described in this article was viewed through two-power binoculars.

It is this reduced distance that goes hand in hand with reduced depth. As in the case of the anaglyph, magnification increases retinal disparities only in proportion to the distance-reduction, not to its square. This is so because disparities increase only due to the image size enhancement. Since the object remains at its actual distance, there is no change in the directions from which the eyes view the object, and the shapes of the projections of the depth-intervals are not altered. Were the distance to the object actually shortened, disparities would increase with the square of the distance-reduction. Stereoscopic depth-constancy largely compensates for this increase in disparity by causing the perceived depth that corresponds to a given disparity to increase nearly with the square of the distance. Inasmuch as the same compensation takes place when magnification increases disparity only in proportion to the distance reduction itself, perceived objects should lose depth under magnification. Ideally, the contraction of perceived depth should be proportional to the reduction in distance.

WHY THE MODIFICATION OF STEREOSCOPIC DEPTH-PERCEPTION IS SO RAPID

By HANS WALLACH and EILEEN B. KARSH, Swarthmore College

Experiments have been reported in which extraordinarly rapid perceptual learning was obtained. By viewing a solid object in rotation through a telestereoscope (an instrument that enhances retinal disparity), the relation between disparity and perceived depth can be changed in a matter of minutes. This modification of stereoscopic perception of depth manifests itself in purely perceptual experiences, and it can be measured by having *O* make estimates of depth before and after the period of exposure to the experimental conditions.[1] The effect appears to be due to the discrepancy between the two sets of cues—those for the kinetic depth-effect and those for retinal disparity.[2] Since the relation between disparity and experienced depth can be modified by what appears to be a learning process, it seems likely that the normal relation between disparity and depth, which *O* brings to the experiment, is also a learned one. For the perception of depth to be veridical, this relation must change during the years of growth, as interocular distance increases, and remain the same afterward. Our *O*s' interocular distances had been constant for a number of years when they participated in the experiment, and the particular relation between disparity and depth that was in existence when our training-procedure changed it had a learning period of several years in back of it. One may then ask why it is that our experimental procedure can rapidly change the quantitative aspect of a perceptual process which has presumably been learned years ago and has been practiced ever since. This paper explores two possible reasons for the rapid learning that is obtained in our experiments.

One explanation makes the assumption that our experimental procedure affords very concentrated learning compared to the learning that takes place in ordinary life. When *O* views the rotating object through the telestereocope, every turn of the object by, say 90° gives rise to a discrepancy between the potential result of the kinetic depth-effect and the enhanced stereoscopic depth that is caused by the telestereoscope. With the

---

* Received for prior publication May 17, 1963. This work was supported by a grant from the National Science Foundation.

[1] Hans Wallach, M. E. Moore, and Linda Davidson, Modification of stereoscopic depth-perception, this JOURNAL, 76, 1963, 191-204.

[2] Hans Wallach and E. B. Karsh, The modification of stereoscopic depth-perception and the kinetic depth-effect, this JOURNAL, 76, 1963, 429-435.

object rotating at approximately 10 r.p.m., the 10-min. training period is packed with occasions which can contribute to the learning of the new relation between disparity and depth. Compared to this, the learning occasions that offer themselves under the conditions one encounters in ordinary life may indeed occur much more rarely, which could account for the fact that a process that undoubtedly has been established for a long time and for which there have since been many learning occasions can be modified in a matter of minutes.

The second explanation assumes that the learning product on which the quantitative aspect of stereoscopic depth-perception is based is not very durable and that in the course of time it has frequently to be reacquired. The training for normal depth-perception would be cumulative only in a limited way and its effect less potent than its long duration would suggest.

EXPERIMENT I A

We tested for the first explanation, which hinges on the temporal concentration of learning occasions in the experimental training, by giving normal depth-perception the same presumed advantage. Prior to the usual training procedure, during which *O* saw the rotating object with artificially enhanced disparity, he observed the rotating object under normal viewing conditions also for a period of 10 min. Such pretraining provides the same kind of training in support of the *normal* relation between disparity and depth which the subsequent telestereoscopic presentation provides for an adaptation to the experimental conditions. Should we find that this pretraining of normal depth-perception strongly diminishes the effect of the telestereoscopic training, the first explanation would be confirmed.

Method. Training and tests proceeded in essentially the same fashion as previously described.[3] The telestereoscope increased the interocular distance by 7.6 cm. For the wire-form, we chose the irregular pyramid previously used. It was four-sided with a nearly square base (the diagonal was 8.5 cm., and the distance from the apex to the base was 3.8 cm.). Presumably this form does not provide perspective-cues for depth when its base occupies the frontal-parallel plane and the apex points toward *O*, the position in which depth-estimates were obtained. It was placed 32.5 in. from *O*'s eyes, with a checkered tablecloth which covered a horizontal platform 16 cm. below the center of the pyramid providing perspective-cues for distance. This tablecloth, stretching from the vicinity of *O* toward the wire-form and 20 in. beyond to a white background-screen, was visible to one eye only. The speed of rotation during the training-period was 10 r.p.m. Depth-estimates were given by adjusting a measuring rod which could be extended to nearly twice its shortest length. *O* held the rod in his hands and looked down at it while using it to represent perceived depth. To make it possible to vary the length of the rods that were given to *O* before and after training, four rods were used that differed from each other in basic length in steps of 1 cm. As in previous work, estimates of size in the frontal-parallel plane were obtained together with the depth-estimates to make sure that the training period did not produce a decrease in total size which might account for part of the measured change in depth. In the case of the pyramid, estimates of the diagonal of the base were obtained.

---

[3] Wallach, Moore, and Davidson, *op. cit.*

Two groups of 20 *O*s were used. One was given the pretraining, which consisted in observing a rotating wire-pyramid under normal viewing conditions, followed by the usual training with the telestereoscope. For the other—the control group—the pretraining was omitted. To make the wire-pyramid presented in pretraining (where it was given with normal disparities) look as much as possible like the pyramid seen through the telestereoscope, the former was larger in its distance from apex to base. While the bases of the two pyramids were identical, the pyramid used for pretraining had a distance from apex to base of 6.2 cm.

The *O*s were selected for good stereoscopic depth-perception and for consistency in their estimates. The following criteria were used: *O*s were eliminated (1) whose initial depth-estimates were more than 0.5 cm. smaller than their estimates of the diagonal; (2) who failed to see the pyramid distort when seen in rotation through the telestereoscope; and (3) whose initial and final estimates of the diagonal of the base differed by more than 0.7 cm.

TABLE I

MEAN ESTIMATES OF SIZE AND DEPTH (IN CM.) FOR TWO INCREMENTS IN INTEROCULAR DISTANCE (EXPERIMENTS I A AND I B)

Increment in interocular distance	Group	Size			Depth			
		initial	final	change	initial	second	final	decrement
7.6	Pretraining	6.325	6.465	0.140	6.690	6.530	5.880	0.81
	Control	6.450	6.510	0.060	6.595	—	5.315	1.28
12.8	Pretraining	6.255	6.305	0.050	7.495	7.055	6.165	1.33
	Control	6.345	6.255	−0.090	7.425	—	5.525	1.90

In detail, the procedure was as follows: All the *O*s were first shown the deeper of the two pyramids, and the particular distances for which they would have to give estimates were pointed out to them. Then they looked through the telestereoscope for 1 min. at the shallower pyramid, which was stationary with its apex pointing toward *O*. A size-and a depth-estimate were then made, in that order. *O* was instructed to base his estimates solely on the appearance of the figure at the time the judgment was given. After the initial depth-estimates, the *O*s of the pretraining group watched the deeper pyramid in rotation under normal viewing conditions, and then made a second depth-estimate of the shallow pyramid seen through the telestereoscope. These two steps were omitted in the case of the *O*s of the control group. Then all the *O*s had the usual training period during which they observed through the telestereoscope the shallow pyramid in rotation, and the experiment ended with the final estimates of depth and size, in that order.

Results. The results are given in the first two rows of Table I. They show a trend in line with the explanation which this experiment tests.

The effect of the usual training period on perceived depth was indeed diminished for *O*s in the pretraining group, who showed a depth-decrement of 0.81 cm., while the mean depth-decrement for the control group was 1.28 cm.; this difference was reliable ($t = 2.58$, $p < 0.02$). The effect of pretraining on the modification of stereoscopic preception of depth was not, however, large enough to establish temporal concentration of learning occasions as the only explanation of the rapid learning produced by our usual training period. To do so, the pretraining should have had a major effect on the depth-decrement, which it did not.

EXPERIMENT I B

Experiment I B, a replication of Experiment I A with a greater increase in interocular distance, served a double purpose. First, we wished to see whether pretraining with the normal relation between disparity and depth would produce a larger change in the usual effect of training when that effect itself became greater. This we hoped to achieve by increasing further the disparity produced by the telestereoscope. Secondly, we wished to find out whether the usual effect of training would indeed become greater when disparity was increased, and, if it did, whether this alteration of our basic procedure would be preferable from a technical point of view.

Method. The arrangement of the mirrors of the telestereoscope was changed to increase each *O*'s interocular distance by 12.8 cm., which, assuming that the mean interocular distance of our *O*s was 6.4 cm., would increase interocular distance by a factor of three. In the previous experiment, interocular distance was increased by 7.6 cm., or, on the average, by a factor of 2.2. To avoid any concomitant enhancement of retinal disparity which, in the case of some *O*s, might result in failure of some part of the wire-form to be binocularly fused, we compensated for the enhancement in interocular distance by increasing the distance from *O* of the form. Since disparity decreases with the square of the distance of the object from the eyes,[4] a moderate increase of that distance from 32.5 cm. to 42.5 cm. was sufficient to achieve such a compensation.

Results. The results (last two rows in Table I) reveal the expected increase in the effect of training; the control group showed a depth-decrement of 1.90 cm., as compared with the decrement of 1.28 cm. in Experiment I A. An additional increase in the interocular distance of 66.6% caused an increase in the effect of training of 48%. The effect of pretraining on training did not, however, increase proportionally with the training effect. Whereas in Experiment I A the depth-decrement after pretraining was 37% less than in the control experiment, the comparable loss in Experiment I B was only 30%. An analysis of the combined results of the two experiments showed a reliable effect of pretraining in reducing depth-loss ($F = 7.90; df = 1, 76; p < 0.01$), but the difference between the two experiments was not reliable.

---

[4] C. H. Graham, Visual perception, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 888.

To be sure, the enhancement of the average interocular distance from 2.2 times to 3 produced a greater training effect, and the difference was reliable ($F = 9.49; df = 1, 76; p < 0.01$), but this advantage was offset by a great increase in variability. While the *SD* of the initial estimate of depth in the case of the smaller interocular distance was 0.93 cm., it amounted to 1.63 cm. for the larger interocular distance, and this difference was significant ($F = 3.05; df = 59, 79; p < 0.01$). Consequently, the loss in depth-decrement due to pretraining in Experiment I B was larger but less reliable.

The larger variability connected with the greater interocular distance is also apparent when the differences between the initial and the second depth-estimates are compard. With only the pretraining period separating the two estimates, there is no reason to expect a systematic change in depth. While for experiment I A the two mean estimates differ only slightly and not significantly (0.16 cm., $t = 0.96$), the results for Experiment I B show a sizable decrement of 0.44 cm., a difference which was not quite significant ($t = 1.86, p < 0.09$). A difference of similar magnitude (0.47 cm.) obtained in Experiment I A between the two depth-decrements was significant, although it pertained to different groups of *O*s It is clear that the tripled interocular distance represents no technical advantage.

EXPERIMENT II

The second explanation for the rapidity with which a modification of stereoscopic depth perception can be achieved assumes a limited durability of the learning product responsible for the quantitative aspect of normal stereoscopic vision. Such limited durability could be demonstrated by showing that a period of disuse makes normal stereoscopic depth-perception less potent and hence less resistant to alteration. We proceeded by restricting *O*s to monocular vision for 24 hr. before the experiment. Originally, we attempted to show that wearing an eye-patch for a day prior to the usual training procedure would enhance the training effect. Such a result would have been an indication of diminished potency of normal depth-perception after a period of disuse and would have served as a demonstration of the specific relation—assumed in our explanation—between the modification of depth-perception and the depth-perception which *O* brings to the experiment. Pilot studies showed that 24-hr. disuse had more serious consequences for stereoscopic vision than was anticipated and that, as a result, our original plan would not succeed. After wearing an eye-patch for 24 hr., *O*s noticed spontaneous changes in depth of a binocularly viewed wire-form. The object would appear alternately deeper and shallower, and, perhaps as a consequence, mean depth-estimates were considerably smaller than normal. This being so, it is not surprising that a subsequent training procedure, while it did result in sizable training effect (that is, a *further* depth-decrement) did not by itself produce the postulated *enhanced* decrement. Accordingly, we altered the goal of the experiment to a mere demonstration of the deterioration of stereoscopic vision which is apparently produced by disuse.

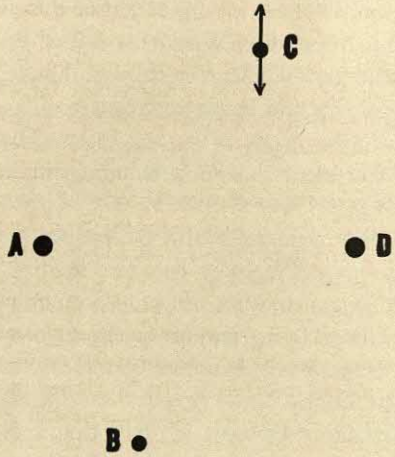


FIG. 1. ARRANGEMENT OF THE FOUR RODS USED FOR DEPTH-INTERVAL MATCHING IN EXPERIMENT II

Method. Two procedures were employed to obtain quantitative manifestations of this deterioration: *O* estimated the depth of a wire-form before wearing the eye-patch and afterward; two depth-intervals were repeatedly matched before and after disuse of binocular vision and the *SD* of each set of matches was determined.

Depth-estimates were given for the deeper pyramid described above. It was placed 42.5 cm. from *O*'s eyes with its apex pointing toward *O*, and it was viewed directly against a white screen. As in Experiment I A the table-surface was raised and covered with a pattern providing perspective-cues to distance. The arrangement for matching two depth-intervals consisted of four vertical rods, 12 in. long, placed according to the scale drawing of the ground plan shown in. Fig. 1. The centers of the Rods *A* and *D* were 4 in. apart and occupied *O*'s frontal-parallel plane 32.5 from *O*'s eyes. Rods *A, B,* and *D* were in fixed position; Rod *C* could be moved toward and away from *O* by a rack-and-pinion arrangement. Placed 5 in. in front of the Rods *A* and *D* was a screen with an aperture, 8 in. wide and 6 in. high, through which *O* viewed them, black against a well-illuminated background. *O*'s task was to match the depth between Rods *B* and *A* with the depth between Rods *C* and *D*. The depth-interval between Rods *A* and *B* (2.5 in.) was so chosen that all *O*s would see both rods fused when they looked at one of them. This was, however, not the case for the Rods *B* and *C* when the latter was in the matching position, that is, as much behind the plane of Rod *A* and *D* as Rod *B* was in front. Each *O* made 10 individual matches and the *SD* of these matches was computed for each *O*. Initially, Rod *C* was set either 2 in. in front or 2 in. behind the matching position. *E* slowly moved it toward the matching position. *O* was permitted to stop this motion to make leisurely comparisons and to ask for a shift in the opposite direction when the point at which the depth-intervals appeared equal seemed to have been passed. We deliberately used this flexible procedure rather than a rigid psychophysical method to keep it adaptable to the difficulties in depth-perception which *O*s have after the period of monocular vision.

Before they were selected, the *O*s were tested for good stereoscopic depth-perception. They were shown a wire-form derived from an octahedron by elongating one of its horizontal axes and asked to give a height- and a depth-estimate, the latter along the elongated axis which nearly fell into *O*'s sagittal plane. This wire-form was 8.8 cm. high and 11.1 cm. deep. Only *O*s whose estimates of depth were 1 cm. greater than their height-estimates were used. The second test consisted of a depth-estimate of the pyramid used in this experiment. To be chosen for the experiment, *O*'s estimate had to be at least 5.5 cm. *O*s who found it difficult to use either eye in monocular vision were eliminated. These were rigorous criteria; of the 60 undergraduates tested, only 20 were used in the experiment.

The sequence in the experiment proper was as follows: *O* made one estimate of the diagonal of the base of the pyramid, and one of its depth. Then he made a group of 10 depth-matches with the vertical bars. One day before the scheduled time for the second test, *O* put on an eye-patch, switched it from one eye to the other every 2–3 hr., and wore it during all his waking hours while he engaged in normal activities. Upon *O*'s return to the laboratory for the second test, the eye-patch was removed, and *O* immediately made a depth-estimate for the pyramid, followed by an estimate of the diagonal of the base. After this, he made another set of 10 depth-matches with the vertical bars.

To get a first idea of the persistence of the deterioration, another set of estimates for the pyramid was obtained from half the *O*s after the matching, which took approximately 7 min. Then the *O*s had to look steadily at the pyramid for 5 min. and make another estimate of depth and of the diagonal. After a rest-period of 1 min., during which the *O*s sat with their eyes closed, they viewed the pyramid again for 5 min. and thereafter made a final set of estimates.

Results. There was a strong decline in the apparent depth of the pyramid after the 24-hr. period of monocular vision. The initial mean depth-estimate was 6.77 cm., while, after wearing the eye-patch, the mean was 5.36 cm. The depth-decrement, which amounted to 20.9% of the original depth was highly significant ($t = 9.19$, $p < 0.001$). There was also a small decrement in size; the mean estimate of the diagonal of the base decreased from 7.26 cm. before to 7.07 cm. after the monocular period, but the difference did not approach significance ($t = 1.64$, $p < 0.20$). The matches of two depth-intervals showed a sizeable increase in variability after the period of monocular vision. The mean *SD* rose from 0.50 cm. before to 0.71 cm. after disuse of stereoscopic depth-perception; this increase was significant beyond the 1% level ($t = 3.00$).

The 10 *O*s used in studying the persistence of the deterioration showed an initial mean decrease of 1.62 cm. in the depth of the pyramid. After the matching test, the mean depth of the pyramid had increased by 0.56 cm.; after the first 5-min. observation of the stationary pyramid, its depth had increased by an additional 0.47 cm., and another 5-min. observation-period produced a further increase of 0.27 cm. At that time, approximately 20 min. after removal of the eye-patch, all but 0.32 cm. of the original depth-decrement had been recovered.

SUMMARY AND CONCLUSION

Two possible explanations for the rapidity with which a modification of stereoscopic depth-perception can be achieved have been tested. One of them assumes that veridical stereoscopic vision is based on a somewhat unstable learning product which must continuously be reacquired on the learning occasions which occur in normal use of binocular vision. We found evidence for a deterioration of stereoscopic depth-perception after a period of 24 hr. of monocular vision, but we have not yet been able to obtain proof of a connection between the striking effect of disuse and the effectiveness of our training procedure.

The other explanation assumes that our training procedure is so effective in changing the product of previous learning in stereoscopic vision which *O* brings to the experiment because a high concentration of learning occasions is provided by our procedure as compared with the frequency with which ordinary life affords them. This explanation was tested by using the same procedure which had been employed to produce the modification of stereoscopic vision to strengthen *normal* stereoscopic perception before the modification was attempted: the *O*s observed a rotating wire-form in direct view for 10 min. before they observed it through the telestereoscope. This experience did result in a diminished modification of stereoscopic vision, but the effect was not large enough to establish high concentration of learning occasions during the rotation as the sole explanation for the rapidity with which a modification of depth-perception can be obtained. To a great extent, this rapidity seems to depend on an inherent modifiability of stereoscopic depth-perception, and our evidence for its fast deterioration supports this view.

SEMANTIC SATIATION AND MEANINGFULNESS

By RABINDRANATH KANUNGO and W. E. LAMBERT, McGill University

'Semantic' or 'verbal satiation' refers to a loss of meaning or a reduction in the effectiveness of verbal material following its continued overt repetition or its prolonged visual fixation. Several attempts have been made to observe and index the effect of verbal satiation.

Smith and Raygor measured word-association and found that after prolonged visual fixation a word loses its common associational response.[1] More recently, Lambert and Jakobovits observed a reduction in the intensity of semantic ratings of words following their continuous overt repetition.[2] They claim to have introduced a "more comprehensive method of measuring the phenomenon" by using scales of the semantic differential as their measuring instruments. In their study, they interpreted the phenomenon as a cognitive form of reactive inhibition.

In a later study, Kanungo, Lambert, and Mauer advanced an alternative interpretation of verbal satiation, considering the phenomenon in terms of principles of associative learning.[3] It was suggested that with more and more repetition of a word, an increasing number of its *m*-components (*i.e.* its repertory of related verbal associates) will gradually drop out. This reduction in associative potential was thought to be possibly attributable to the fact that through continuous repetition a word tends to evoke itself increasingly as the most probable associate. Thus repeating "table, table, table . . ." will increase the tendency for 'table' to evoke 'table' as an associative response. This interpretation is based on Noble's "theoretical-experimental analysis of attributes of meaning in verbal stimulus-material."[4] Noble has tried to develop his analysis of meaning within the framework of contemporary theory of behavior. He suggests that, "Since theoretically the number of responses is proportional to the number of supraliminal E's, frequency of response is proposed as a rational index of stimulus-meaning (*m*). Therefore, the appropriate association-value for the present analysis would be denoted by the central tendency of the frequency-distribution of continued associations given by Ss per unit-time."[5]

---

* Received for publication October 18, 1961. This research was supported in part by the Canadian Defence Research Board, Grant number 9401-10.

[1] D. E. P. Smith, and A. L. Raygor, Verbal satiation and personality, *J. abnorm. soc. Psychol.*, 52, 1956, 323-326.

[2] W. E. Lambert, and L. A. Jakobovits, Verbal satiation and changes in the intensity of meaning, *J. exp. Psychol.*, 60, 1960, 376-383.

[3] Rabindranath Kanungo, W. E. Lambert, and S. M. Mauer, Sematic satiation and paired-associate learning, *J. exper. Psychol.*, 64, 1962, 600-607.

[4] C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421-430.

[5] Noble, *op. cit.*, 424.

The purpose of the present study is to extend this index of meaningfulness to the measurement of semantic satiation and to explore the relation between the process of semantic satiation and verbal association.

Method. Subjects. Thirty-two English-speaking cadets of the Royal Canadian Air Force enrolled in a six-week training course at a base near Montreal, were asked to volunteer for the experiment by their instructor. Each *S* was tested individually at the training base during regular work hours.

Materials and procedures. Sixteen two-syllable nouns, chosen from Noble's list, were used for the present study on the basis of their high *m*-values. Four different sets of four words each were included. These will be referred to as Sets A, B, C and D. Care was taken to equate sets on the basis of their *m*-values as given by Noble.[6] Each of these words was printed on a separate 5 × 3-in card.

The *Ss* were tested twice, with 24 hr. intervening between the first and the second testing. In the first testing, *m*-values for each set of words were obtained. In the present study, the *m*-value of a set is defined as the grand mean number of 'relevant' responses given within a 60-sec. period by all *Ss*. The procedure was similar to that used by Noble.[7] Exactly the same instructions were used and the stimulus-words of each of the sets were administered to *S* in a test-booklet with answer-sheets attached. The booklet contained one stimulus-word on each page and, to reduce *S*'s tendency toward free association, the word was reproduced on each line so that he would be more likely to return to the stimulus-word before making a new response. *S* was allowed to write associations to each stimulus-word for a period of 60 sec. Immediately preceding the first testing, two words, 'leader' and 'effort,' were administered as practice-words. A counterbalanced order of presentation of words from each set was adopted for which an equal number of *Ss* were randomly assigned to each of the following four orders of presentation: AAAABCCBBCCB, AAAACBBCCBBC, BDDEBDDBAAAA, DBBDDBBDAAAA. Thus in the first testing, for all *Ss* the booklet contained words from Sets A and B, but for half of the *Ss* it contained words from Set C and for the other half it contained words from Set D. In the second testing, however, the booklet contained stimulus-words from both Sets A and B, but those *Ss* who were administered C-words in the first testing received D-words in their second booklets, and those who had D-words in their first testing received C-words in their second booklets. Before an *S* started giving associations in the case of words from Set A, he was shown a card with the stimulus-word printed on it for 1 sec. He read the word aloud and was asked to write associations to that word for a period of 60 sec. In case of each of the other words from Sets B and C or D, *S* was asked to look fixedly at the word printed on a card and repeat it aloud at the rate of 2–3 repetitions per sec. for a period of 20 sec., and immediately after to write associations to that word for a period of 60 sec. Thus for purposes of satiation, both the methods of visual fixation and continuous verbal repetition were used.

It will be observed that stimulus-words from Set A served as the retest-control condition in which the same words were administered twice to the same *Ss* and their *m*-values were obtained. Stimulus-words of Set B served as Experimental

---

[6] Noble, *ibid.*, 426.

[7] Noble, *ibid.*, 424.

Condition I in which *m*-values for the same words were obtained from the same *S* both before and after satiation. Words of Sets C and D combined served as Experimental Condition II, in which *m*-values for the same words were obtained both before and after satiation treatment but from different *S*s. This latter procedure eliminates the influence of *S*'s responses to words in the first testing (*i.e.* before satiation) upon his responses to the same word in the second testing.

For scoring purposes, all the responses that *S* wrote were taken into consideration. These were divided into two categories, 'relevant' or acceptable, and 'irrelevant' or unacceptable. This distinction of relevant responses was made after examining the criteria developed by Noble for relevant and irrelevant responses. Thus, illegible responses, perseverative responses, free or tangential associations, and clang-associations were considered irrelevant responses. Relevant responses, then, defined the *m*-value of any word.

Results and discussion. The sets of words used and their average *m*-values, as determined by Noble, are presented in Table I. Their mean *m*-values as determined in the present study are also presented in Table I. These were calculated for each set after the first testing. It will be observed that, in spite of the small number of *S*s, the mean *m*-values per word in each set are quite close to those of Noble's. Only in case of Set B is the mean *m*-value per word higher, the difference being 1.07 per word.

TABLE I

FOUR SETS OF DISSYLLABIC WORDS (NOUNS) SELECTED FROM NOBLE'S LIST AND MATCHED FOR THEIR MEAN *m*-VALUES

	Mean *m*-value			
	Set A	Set B	Set C	Set D
Words	UNCLE VILLAGE OFFICE MONEY	YOUNGSTER GARMENT JEWEL DINNER	CAPTAIN ZEBRA JELLY WAGON	ZERO INSECT HEAVEN ARMY
(Noble's study)	7.69	7.45	7.45	7.72
(present study)	7.65	8.52	7.59	7.48

An analysis of variance design appropriate for Treatment × Treatments × *S*s was chosen from Lindquist,[8] since each of the 32 *S*s was tested twice under three different conditions without an effect of order. It may be recalled here that, to avoid any possible effect of sequence of presentation, a counterbalanced design was adopted. Thus there are two main treatment-classifications. One is the first and the second tests, and the other is the three experimental and control conditions. In the case of relevant associations, the difference between Conditions is significant ($F = 27.96$, $df = 2/62$, $p < 0.01$) as is the interaction of Tests and Conditions ($F = 14.27$, $df = 2/62$, $p < 0.01$). But in case of irrelevant associations, both main effects, as well as their interaction, attain the 1% significance. The F-ratios are 9.80, 9.47, and 19.89 for Conditions, Tests and the interaction, with $df = 2/62$, $1/62$ and $2/62$, respectively.

---

[8] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1956.

Comparison between mean number of associations in the first and in the second tests, for each of the three conditions, is shown in Table II. These comparisons were carried out for both the relevant and the irrelevant associations. Here the mean number of responses in any given condition is calculated by dividing the total number of written associations (relevant or irrelevant, as the case may be) given to all the stimulus-words administered in that condition by the total number of *S*s. For testing the significance level of mean differences, the critical difference corresponding to 1% level was calculated.[9] Any mean difference that exceeds 2.89 for relevant responses and 1.36 for irrelevant responses is considered significant at the 1% level. Results in Table II reveal that, both for the Control Condition and for Experimental Condition I, there is no significant difference between the mean number of irrelevant responses on the first and on the second tests, but that there is a significant increase ($p < 0.01$) in the mean number of relevant responses in the second test. In Experimental Condition II, however, the mean number of relevant responses in the second test drops significantly and the mean number of irrelevant associations shows a significant increase, both significant at the 1% level.

TABLE II

COMPARISON OF MEAN NUMBER OF ASSOCIATIONS IN THE FIRST AND THE SECOND TESTING

Conditions	Nature of associations	First testing		Second testing		Mean diff.
		M	*SD*	*M*	*SD*	
Control	relevant	30.59	8.38	33.87	10.18	3.28*
	irrelevant	2.34	3.41	2.47	3.27	0.13
Experimental I	relevant	34.06	9.64	37.06	10.15	3.00*
	irrelevant	1.69	2.62	2.38	2.69	0.69
Experimental II	relevant	30.17	8.09	26.19	7.68	3.98*
	irrelevant	1.34	2.55	5.69	4.77	4.35*

* Significant at 1% level.

---

[9] E. F. Lindquist, *op. cit.*

Table III presents comparisons of the conditions, taken two at a time, for their mean number of relevant and irrelevant responses considered separately. This was done both for the first and second tests. In the first test, there is no difference between the three conditions with respect to their mean number of irrelevant responses. In the second testing, only Experimental Condition II shows a significant increase ($p < 0.01$) in its mean number of irrelevant responses as compared to either the Control Condition or the Experimental Condition I. No difference in the mean number of irrelevant responses could be observed between the Control Condition and Experimental Condition I in the second test.

With respect to relevant responses, Experimental Condition I has a significantly higher mean than either the Control Condition or Experimental Condition II in both the first and the second tests. When Experimental Condition II is, however, compared with the Control Condition, it is seen that though their mean number of relevant responses does not differ in the first test, they do differ significantly ($p < 0.01$) in the second test.

TABLE III

MEAN DIFFERENCES BETWEEN CONDITIONS IN THE NUMBER OF ASSOCIATIONS

Conditions	First testing		Second testing	
	relevant	irrelevant	relevant	irrelevant
Control vs. Experimental I	3.47*	0.67	3.19*	0.09
Control vs. Experimental II	0.44	1.00	7.69*	3.22*
Experimental I vs. Experimental II	3.91*	0.35	10.88*	3.31*

* Significant at 1% lelvel.

The results for relevant responses clearly show that both in the Control Condition and Experimental Condition I, there is a significant mean increase in the second test. In Experimental Condition II, instead of an increase, there is, however, a significant mean drop in the second test (see Table II). This is why the interaction of Tests with Conditions is significant and also why the main effect of Tests does not attain significance. Further, since in the second test there is a significant mean drop in relevant responses for Experimental Condition II, and a significant increase in relevant responses for the other two conditions (see Table II), the initial mean difference in the first test between Experimental Condition II and either of the other two conditions becomes greater in the second test. The initial mean difference in relevant responses between Experimental Condition I and the Control Condition remains fairly constant, however, for both first and second tests (see Table III). Thus, the only condition that behaved differently from either of the two conditions mentioned above is Experimental Condition II.

With respect to the irrelevant responses, only in the case of Experimental Condition II is there a significant mean increase in the second test (see Table II). This is again the reason why this condition differs from the other two, not in the first test but in the second test (see Table III). Of course, the interaction of Tests with Conditions is also significant.

These results suggest that two of the conditions, Experimental Condition I and the Control Condition, were alike in the second test with respect to both relevant and irrelevant associations. In both conditions there is a significant increase in their respective mean relevant responses and almost no change in their respective mean irrelevant responses. Even the initial amount of difference in their mean number of relevant responses in the first test is almost similar to the difference obtained in the second test (Table III). Thus, the only condition that behaved differently from either of the two conditions mentioned above is Experimental Condition II. The *Ss* under this condition showed a significant mean drop in relevant responses and a significant mean increase of irrelevant responses in the second test.

In the Control Condition the same words were administered a second time after 24 hr. The results indicate that there is a significant mean increase of relevant responses and no change in irrelevant responses. Since the same words were administered during the second test, the memory of the previously written responses may account for such an increase in relevant responses. In Experimental Condition I, satiation was given to the same words to which the *Ss* gave written responses in the first test. Thus it seems that in this condition also the memory of the previously written responses may have facilitated the production of relevant responses during the second test. A facilitation of memory for relevant responses would have given little scope for satiation to show an effect. This hypothesis receives substantive support from the data when the relevant responses of the second testing are examined more closely. For the Control Condition, 45% of the responses, and for Experimental Condition I, 51% of the responses are direct reproductions from the first test.

The mean number of relevant responses in Experimental Condition II dropped, and the mean number of irrelevant responses increased. This condition was especially designed to avoid any kind of interaction of memory with satiation. We have already seen that such an interaction is very likely to reduce the effect of satiation. The results obtained with Experimental Condition II, therefore, can only be attributed to satiation of novel words in the second test. It is possible that the results of Experimental Condition II can be attributed to appearance of novel words in the second test rather than to satiation, *per se*. To determine if the appearance of new words in the second test tends to reduce the number of relevant responses and increase the irrelevant ones, 16 college freshmen were administered 4 stimulus-words in the first test, and after a two-day interval were administered 4 different words in a second test. The same instructions and procedure as described earlier were used. The eight words were taken from Sets A and B as presented in Table I. To avoid any effect of specific words, all 8 words were randomly assigned to the first and the second test for each *S*. The results are presented in Table IV. It is apparent that with the appearance of new words in the second test, there is a trend toward increase in the mean number of relevant responses and a decrease in the mean number of irrelevant responses (in each case $p < 0.10$). These results can only be accounted for in terms of a transfer of practice from the first to the second test. Since the results obtained for Experimental Condition II are just the opposite of what is presented in Table IV, the hypothesis that mere appearance of novel words in the second test can account for the results presented earlier for Experimental Condition II is ruled out. Thus the reduction in the mean number of relevant responses in the second test in Experimental Condition II suggests that satiation results in a blocking of relevant responses. But the reason for an increase in irrelevant responses during the second testing under Experimental Condition II, and not under the other two conditions, lies in some factors other than satiation. The authors are inclined to attribute the increase in irrelevant responses to an 'instructional set' rather than to satiation *per se*. *S* was given the instruction that he was taking a *test* to determine how many words he could think of and write down in a short time after seeing a key-word. When relevant re-

TABLE VI

COMPARISONS OF MEAN NUMBERS OF RESPONSES WHEN NOVEL WORDS APPEARED IN THE SECOND TEST

Nature of associations	First test		Second test		t	p
	Mean	*SD*	Mean	*SD*		
Relevant	30.56	5.13	33.50	7.37	1.85	$<.10$
Irrelevant	3.69	2.17	2.63	1.57	2.12	$<.10$

sponses to the key-word are blocked in Experimental Condition II, *S* may attempt to fill in as many spaces on a page during the specified time and even turn to a repertory of irrelevant responses to do so. An examination of irrelevant associations for this condition reveals an abundance of free and tangential associations. In Experimental Condition I, however, the blocking effect of satiation did not show up, possibly because of the facilitating effect of memory. In this case, *S*s had no need to revert to irrelevant associations.

From another perspective, it is an interesting possibility that stimulus-words which have been satiated may become more vague and less meaningful so that *S*s may actually believe that what are judged as irrelevant responses are in fact relevant for stimulus-words whose meanings have been modified.

SUMMARY

The present study employs Noble's concept of meaningfulness (*m*) as an index of loss of meaning attributable to verbal or semantic satiation, and explores the relation between semantic satiation and verbal association. Sixteen two-syllable nouns were chosen from Noble's lists, and four different sets of four words each were prepared. Each *S* was tested twice with a 24-hr. interval between tests under three different conditions. Words from one set were used in a test-retest control condition in which the same words were administered twice to the same *S*s and their *m*-values obtained each time. Words from a second set were used in Experimental Condition I in which *m*-values for the same words were obtained from the same *S*s both before and after satiation-treatment. Words from two other sets were used in Experimental Condition II in which *m*-values for the same words were obtained but from different *S*s. Both the relevant responses and irrelevant responses were taken into account in the analysis.

The results showed that for both the Control Condition and Experimental Condition I there was a significant mean increase of relevant responses and no change in irrelevant responses in the second test. Since in both these conditions the same words appeared in the two tests for each *S*, the results were attributed to the facilitative effects of memory. In Experimental Condition II a significant mean drop in relevant responses and a significant mean increase in irrelevant responses were found. It was suggested that satiation leads to a blocking of relevant associations and simultaneously modifies irrelevant associations.

THE MODIFICATION OF STEREOSCOPIC DEPTH-PERCEPTION AND THE KINETIC DEPTH-EFFECT

By HANS WALLACH and EILEEN B. KARSH, Swarthmore College

An earlier article described investigations of a novel process of perceptual learning, the modification of stereoscopic depth-perception by means of slow rotation.[1] In viewing a rotating three-dimensional wire-form through a telestereoscope, the normal relation between disparity and apparent depth was modified in a matter of minutes. A 'training period' of 10 min. produced a reduction of perceived stereoscopic depth by as much as 20%. Evidence for such a change was obtained by having *O*s make estimates of depth before and after the training period, or by obtaining descriptions of a rotating wire-form seen directly; the form appears to distort as it turns as a result of a shortened depth-dimension. This training effect transfers to different wire-forms in diminished amount, and it dissipates with mere passage of time. The effect grows in steady fashion with the length of the training period not only for the group as a whole but also for the individual *O*s.

Among the problems that remained unanswered was in what way rotation causes the training effect. We see two possible answers:

Hypothesis 1. Rotation makes the enhanced depth caused by the telestereoscope immediately noticeable through the ceaseless distortion of the turning wire-form. Any cross section of the wire-form that is parallel to the axis of rotation gains in width as it turns from a frontal-parallel position to one in depth where the enhanced stereoscopic depth affects it, and it becomes narrower again as it returns to a frontal position where image-shape again determines its width. A tendency toward invariant form might be postulated to account for the training effect which, by diminishing the apparent depth that corresponds to the given retinal disparity, diminishes also the form-distortions that result from the enhanced retinal disparities.

Hypothesis 2. By providing the conditions of stimulation for the kinetic depth-effect (*KDE*) which also mediates the perception of three-dimensional form and hence of depth, the rotation of the wire-form in our training situation produces a perceptual discrepancy, and the training effect may be based on a tendency to reduce this discrepancy. When there are two different conditions of stimulation

---

* Received for prior publication March 12, 1963. This work was supported by a grant from the National Science Foundation.

[1] Hans Wallach, M. Moore, and L. Davidson, Modification of stereoscopic depth perception, this JOURNAL, 76, 1963, 191-204.

that cause or affect the same perceptual property, the two resultant perceptual processes operate in such a way that, under normal conditions, they produce the same value of that perceptual property—depth in our case. The two sets of stimulus-conditions considered here are (a) retinal disparity and (b) deformations of the retinal projection of the wire-form due to its rotation, which, owing to the *KDE*, also give rise to the perceived object's depth. For an *O* with adequate stereoscopic depth-perception, the two simultaneously operating perceptual processes will normally produce the same apparent depth; but, when one set of cues is manipulated to cause the result of one perceptual process to differ from the other, as happens in our training period where stereoscopic depth is enhanced, a perceptual discrepancy comes into existence. A tendency to reduce this discrepancy through the modification of one of the two perceptual processes may be postulated as the cause of the training effect.

How can one hope to obtain evidence favoring one of these explanations over the other? It is important to note that Hypothesis 1 refers only to the objective fact of rotation, whereas Hypothesis 2 is concerned essentially with the psychological consequence of rotation—the *KDE*—which causes depth-perception in its own right. According to Hypothesis 1, rotation 'exposes' the telestereoscopic effect by causing axial cross-sections to change in width as they successively occupy a frontal-parallel and a sagittal plane. According to Hypothesis 2, there is a change in stereoscopic depth-perception because, with the involvement of the *KDE*, two different processes of depth-perception pertaining to the same form take place simultaneously and cause conflicting results. Proof of involvement of the *KDE*, therefore, would be the answer to our problem, and such proof could be obtained if some characteristic of the *KDE*, not shared by mere rotation, could be shown to influence our training effect. Such a characteristic emerged from a study of the *KDE* by Wallach and O'Connell.[2] One of the conclusions of the study was that not just any deforming retinal image produced by a rotating solid will cause a reliable *KDE*. Only images displaying edges that change simultaneously in length and orientation will produce the *KDE* regularly, unaided by set or expectation. Images whose edges are only parallel and perpendicular to the solid's axis of rotation, and which therefore do not undergo changes in orientation, will cause, without the aid of set, the experience of a turning three-dimensional form only for some *O*s. Some evidence was reported in support of this claim.

Exploratory experiments by Stollnitz done in our laboratory five years ago demonstrated very strong effects of set which emerged as a potent condition influencing whether or not solid forms will be experienced. Immediately preceding experiences with two-dimensional and deforming shapes often seem to be able to prevent the *KDE* from occurring, while a set for seeing three-dimensional forms in rotation will cause such forms to be perceived even when the given pattern displays only contours that change in length alone. The claim that only simultaneous changes in length and orientation of edges will reliably cause the *KDE* can apparently be demonstrated only under set-free conditions. Still, we believe that the claim is concerned with a genuine characteristic of the *KDE* and can, thus, be used as a criterion as to whether the *KDE* as such is involved in our training effect.

---

[2] Hans Wallach and D. N. O'Connell, The kinetic depth effect, *J. exp. Psychol.*, 45, 1953, 205-217.

Our experiment, then, consists essentially in using, in the situation that yields the modification of stereoscopic depth-perception, two wire-forms which differ in whether they display edges that change simultaneously in length and orientation during rotation. If the training effect obtained with the wire-form displaying such edges is larger, a strong argument can be made for the participation of the *KDE* in our training effect. This conclusion, in turn, would support Hypothesis 2. If no difference were obtained, the experiment would be inconclusive. That result would, of course, follow from Hypothesis 1, but it could also occur because, in our training situation, set for the *KDE* to take place was maximal, since binocular disparity also provides cues for depth and rotation and can cause these experiences independently of the *KDE*.

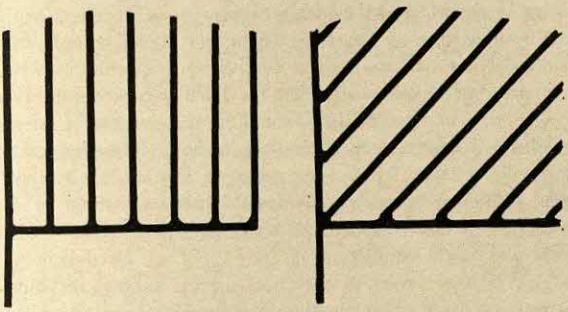


FIG. 1. WIRE-FORMS USED TO TEST HYPOTHESIS 2

Method. We made the wire-forms for this experiment as simple as possible. They consisted of flat arrangements of parallel wires attached to the supporting shaft like flags (Fig. 1). When each figure was turned about the vertical shaft, the flag with the vertical wire arrangement produced a retinal image which deformed in such a way that mainly the length of the supporting horizontal wire and the horizontal distances between the parallel wires changed periodically. For the 'diagonal' flag, however, the oblique wires underwent in addition simultaneous alterations in length and orientation, changing from a 45° tilt when the flag was in a frontal-parallel position to one approaching verticality as the flag approached the sagittal position. The curved upper edges of the flags were meant to minimize perspective cues for depth which the flag's shape might provide. It had previously been shown that a modification of stereoscopic depth perception can be obtained with stationary exposure of orthogonal wire-forms to telestereoscopic viewing; that is, perspective cues provided by the wire-form can, if only in a minor way, have an effect similar to rotation.[3] This fact also made it necessary to use control conditions where the training period consisted in viewing a stationary flag through the telestereoscope, in order to measure any contribution of perspective cues to the training effect obtained with the flags.

---

[3] Wallach, Moore, and Davidson, *op. cit.*, 191-204.

The apparatus was previously described.[4] The version with the checkered table cloth visible to one eye was used. The telestereoscope added 7.6 cm. to *O*'s interocular distance. The flag was 32.5 in. from *O*'s eyes and visible against a white screen 20-in. beyond. It was not fully rotated, but was made to oscillate through an angle of 90° at a rate of 17 c.p.m. between the frontal-parallel position and the sagittal position pointing away from *O*. Depth- and size-estimates were made with the flag fixed in a position 60° from the frontal plane toward the rear. In making size-estimates, *O*s were asked to reproduce the apparent height of the flag by adjusting a brass measuring rod. Six rods of various lengths were used, each of which could be extended to almost twice its shortest length. The *O*s gave depth-estimates by reproducing the apparent length of the flag which in its 60° position was, of course, enhanced by the effect of the telestereoscope.

There were four groups of *O*s; two groups saw their flag oscillate during the training period and two (in the control condition) inspected it, through the telestereoscope, in the same fixed position as for the depth-estimates. For all groups, the training period was preceded by a 1-min. inspection-period, during which *O* looked through the telestereoscope at the flag in the fixed position, and by the first height- and depth-estimates, which were given in that order. The training period itself consisted of two observation-periods of 2-min., separated by a 15-sec. rest-period. The training period was immediately followed by the second depth-estimate and then the second height-estimate.

Altogether 64 *O*s were used, 20 in each of the oscillation-groups and 12 in each of the stationary-exposure groups. They were selected for adequate steroscopic depth-perception and for consistency in their estimates. Stringent criteria were adopted, and 68 *O*s had to be discarded, mostly for defects in stereoscopic vision. The criteria were as follows:

(1) Ability to see all parts of the flag fused simultaneously in spite of the enhanced disparities that the telestereoscope produces.

(2) A first estimate of the flag's length at least 1.5 cm. greater than the first height-estimate.

(3) No greater difference between the two height-estimates than 0.5 cm.

(4) Reports that the flag appears to change in length when its oscillation is observed through the telestereoscope. To make it possible to employ this criterion in connection with the stationary groups, *O*s of those groups, after completion of the procedure assigned to them, were given training with their flag in oscillation.

We believe that the high reliability of the results to be reported is due in part to the careful selection of *O*s.

Results. The results are given in Table I. The diagonal flag in oscillation produced by far the greatest training effect, a decrement in depth of 0.75 cm. which is very reliable ($t = 8.26$; $p < 0.001$). Compared with this, the effect obtained with the vertical flag in oscillation, a decrease of 0.205 cm., is almost negligible, although it is still reliable ($t = 2.15$; $p < 0.05$). As the height-estimates indicate, these depth-decrements cannot be attributed to decreases in total size. The height-estimates after training are on the contrary a bit larger than before ($F = 9.35$; $p < 0.01$) and are about of the same order of magnitude for all groups ($F < 1.0$).

---

[4] *Ibid.*, 191-204.

An analysis of variance was performed on the depth-estimates of all *O*s before and after training. It showed that the training period caused a significant decrease in depth-estimates ($F = 26.56$, $p < 0.001$). The striking result was that the group with the diagonal flag in oscillation, when contrasted with the three other groups, accounted for 99.7% of the sum of squares of the interaction of treatments (the various training conditions) and training effect. Thus the training effect for this group differed reliably from that of the other groups ($F = 20.20$, $p < 0.001$). There were no differences between the group with vertical flag in oscillation and the stationary groups, as shown by a minute residual interaction.

TABLE I

MEAN ESTIMATES OF DEPTH AND HEIGHT (IN CM.) FOR FOUR GROUPS

Group	N	Depth before training	Depth after training	Depth decrement	Height before training	Height after training	Height increment
Diagonal, oscillation	20	5.890	5.140	0.750	4.000	4.045	0.045
Vertical, oscillation	20	6.065	5.860	0.205	3.955	4.020	0.065
Diagonal, stationary	12	5.500	5.400	0.100	3.500	3.658	0.158
Vertical, stationary	12	6.158	6.092	0.066	3.875	3.950	0.075

Thus the significant training effect obtained with the vertical flag is probably not due to oscillation which was absent from the training period of the other two groups. Rather, the small effect obtained with these three groups is likely to be due to a factor they have in common, probably some residual perspective-cue.

There is then strong evidence that the large training effect that, in the case of our flags, is obtained only through training with oscillation is connected with the *KDE*. Only the diagonal flag, which in oscillation produces those deformations of its retinal images that genuinely give rise to the *KDE*, yields the large training effect. The vertical flag, producing no such deformation, fails to do so. That is the result which, we stated, would confirm Hypothesis 2. Hypothesis 1, by contrast, would predict equally strong training effects for both flags in oscillation. By presenting a flag alternately in the frontal position where its length is given by means of its retinal projection and in depth where that length is given largely by means of retinal disparity, oscillation provides the information about the non-veridical stereoscopic viewing conditions, and it does this equally well for both flags. The training effect should therefore be the same for both flags. The fact that a training due to oscillation was obtained only with one of the flags is, therefore, not in agreement with Hypothesis 1.

Discussion. Some readers may favor a third hypothesis that ascribes our training effect to the *experienced* distortions of our rotating wire-forms. Even without the present evidence we consider this a most unlikely explanation. Perceptual processes are generally inaccessible to awareness, and only their results may come to be represented in experience. That an experienced fact, by virtue of being a matter of awareness alone, should be a causal link in some perceptual process, as this hypothesis proposes, is thus rather unlikely. Quite apart from such considerations, there does not appear to be any reason why oscillation of the diagonal flag should lead to experienced distortions and oscillation of the vertical flag should not. To check whether, inexplicably, this did occur, we asked all our *O*s if they noticed a change in length while the flags were oscillating and found that most of them did. It will be remembered that report of distortion during oscillation was one of the criteria for the selection of *O*s. Of all *O*s tested in selecting the two oscillation-groups, only five were eliminated because they did not report changes in length during oscillation; of these, two were tested with the diagonal and three with the vertical flag.

Once it is clear that the *KDE* is involved in our training effect in a decisive manner, the notion of perceptual discrepancy follows. Stereoscopic vision on the one hand and the *KDE* on the other represent two parallel but separate routes to the perception of three-dimensional form and of depth. The former is strictly a matter of binocular stimulation. The latter is based on a kind of stimulation which is totally effective in a single eye. Whereas, in stereoscopic vision, near and far are strictly determined, the *KDE* leaves the sagittal dimension of visual space undetermined: in the absence of other cues for depth, monocularly given rotating wire-forms readily invert, and reverse depth is as frequently perceived as true depth. When, due to the effect of the telestereoscope, retinal disparity misrepresents objective depth, the outcomes of the two perceptual processes will be in disagreement and perceptual discrepancy results.

One might object to the concept of perceptual discrepancy on the ground that it is too far removed from facts, that it is concerned with the hypothetical outcome of hypothetical processes. This is true enough; as stated, the concept does not concern stimulation. Nor is it a matter of perceptual experience; as far as we know, a discrepancy is never experienced as such. The perceptual result is either experience being dominated by one of the competing processes—as seems to be true of our case in which stereoscopic vision dominates—or experience representing a compromise between the expected results of the two processes—Asch and Witkin's findings seem to provide an example of this.[5] The concept is soundly based nevertheless, since our discrepancy is the normal outcome of clear-cut facts of stimulation. With contradictory sets of sensory cues at the root of the perceptual discrepancy, one might, in fact, speak of cue-discrepancy instead. We are perfectly happy with this formulation as long as it remains clear that cue-discrepancy is defined here by the normal results of the ensuing perceptual processes and not in relation to the physical object. That one set of cues represents the objective situation correctly while the other does not is totally irrelevant, unless a third set of cues that also mediates the objective situation is given.

Since there are a number of perceptual properties that are dependent on more than one set of cues, cue-discrepancy may have wide application in perceptual learning. It may, in many cases, replace current explanations of perceptual learning which ascribe an essential role to motor activities. According to this view, the perceptual result of action would provide the information needed to make perception more veridical. Needless to say, such feed-back is not involved in the modification of stereoscopic depth, simply because *O* remains inactive during the training period.

RELEVANCE OF THE RESULTS FOR THE *KDE*

Our demonstration that the kinetic depth-effect *KDE* plays an essential role in the modification of stereoscopic depth-perception is of importance also in connection with the *KDE* itself, for our experiment corroborates the older claim that only those rotating objects will give rise to the *KDE* whose deforming retinal images have edges that change simultaneously in length and orientation. To be sure, this claim was originally supported by an experiment in which these conditions *always* produced rotation in depth, whereas change in length alone had this effect only in the case of one-third of the *O*s; but a different kind of evidence is of great value. It is provided by the result of our experiment in which the diagonal flag gave rise to perceptual learning due to oscillation, while the vertical flag did not. The two flags differed in that the projection of the former provided the stimulus-pattern claimed to be essential in producing a *KDE*, while that of the vertical flag showed only changes in length.

---

[5] S. E. Asch and H. A. Witkin, Studies in space perception: II. Perception of the upright with displaced visual field and body tilted, *J. exp. Psychol.*, 38, 1948, 455-477.

DETERMINANTS OF THE EFFECT OF POSITION IN SERIAL LEARNING

By ROBERT E. BOWMAN and WILLARD R. THURLOW,
University of Wisconsin

The bowed effect of serial position in rote-learning has been demonstrated in numerous experiments and several theories have been advanced to account for it. In general, the effect is explained on the basis of the positions which a fixed order of syllables occupy within a trial over a succession of similar trials. For example, Hull *et al.* postulate inhibitory and excitatory potentials which accumulate in a particular way dependent on the position of the syllable within the trial.[1] Similarly, theories based on proactive and retroactive inhibition, association with position, or forward and backward learning, have been described as if the postulated processes were being turned off and on by the interpolation of the inter-trial interval between each ordered presentation of the serial list.[2] An experiment by Mitchell,[3] and its recent replication by Eysenck,[4] in which the usual effect of serial position was obtained despite the elimination of the inter-trial interval, questions the validity of these theories. A more stringent test is, however, possible.

If, as these theories imply, the processes underlying the effect of serial-position should begin and end with each trial, it should be possible to equalize the effects of these processes on all syllables by systematically changing the position of each syllable within the successive trials while holding constant the serial order of the syllables. For example, if there are 13 syllables in a list, and each syllable is always presented in the same relative position in respect to the other syllables, it is possible to present every syllable at every position during the course of 13 trials. If this were

---

* Received for publication September 20, 1961. This paper is based in part on a thesis submitted by the senior author in partial fulfillment of a Master of Science degree in the Department of Psychology of the University of Wisconsin.

[1] C. L. Hull, C. I. Hovland, R. T. Ross, M. Hall, D. T. Perkins and F. B. Fitch, *Mathematico-Deductive Theory of Rote Learning*, 1940, 21-81.

[2] J. A. McGeoch and A. L. Irion, *The Psychology of Human Learning*, 1952, 125-134.

[3] M. B. Mitchell, The effect of serial position in the continuous memorization of numbers, this JOURNAL, 45, 1933, 493-494.

[4] H. J. Eysenck, Serial position effects in nonsense syllable learning as a function of interlist rest pauses, *Brit. J. Psychol.*, 50, 1959, 360-362.

done and the correct responses were plotted against positions, theory predicts that the effect of serial position would not be observed.

The present experiment tests this prediction. The design of the experiment also makes possible a test for within-trial effects of position, since serial or list-positions are counterbalanced at each position within the various trials.

Method. (1) Apparatus and stimuli. A memory-drum equipped with a shutter was used. Two lists, each of 13 syllables of low associative value, were devised according to the rules by Hilgard.[5] These lists were typed in capital letters on a circular band of white vellum in the following order: List A: LIW, DAJ, YEB, VUR, CAQ, NID, FEP, ZOS, BEX, KAZ, QUN, GIC, MEF; List B. GIZ, SOQ, BUP, YAT, FEX, KOC, RUV, LAH, MIJ, XOW, VEB, QUN, DAK.

(2) Subjects. Five groups of *S*s were used, each composed of 13 undergraduates having no prior experience with experiments using a memory-drum.

(3) Design. Three successive experiments were performed. Experiment I was designed to confirm pilot-studies which indicated that both a serial and a within-trial effect of position could be obtained by counterbalancing serial positions at within-trial positions, and that the latter effect was obtained only from *S*s who reported learning the syllables as a single list. The experimental method of presentation was used with two groups of *S*s. Treatment of the groups differed only in that in Group 1 the serially first syllable of the list was the only formal anchor, whereas in Group 2, the fifth and ninth syllables, as well as the first, were marked to act as anchors.

Since the effect of serial position obtained in Experiment I appeared qualitatively very similar to the classical effect,[6] Experiment II was run for quantitative comparison. In addition, certain procedural changes were instituted to permit a more rigorous interpretation of the results. Again, two groups, Groups 3 and 4, were used. Group 3 was given the experimental treatment; Group 4, the classical treatment in which the serially first syllable of the list was presented first on every trial. The attempt in Experiment I to induce the learning of the serial list in short segments rather than as a unitary list appeared to have failed, since the curves for the two groups (Groups 1 and 2) looked identical. Experiment III was therefore performed with a single group (Group 5) as another attempt to induce such serial learning of clusters.

(4) Procedure. To present each serial position at every within-trial position in the experimental groups, the order of presenting syllables to each *S* was arranged as a 13 by 13 square in which columns represented trials and rows represented within-trial positions. The syllables were presented down each column in order from whichever syllable started that column, and each succeeding column (or trial) began with the fourth next serial position from the starting syllable of the preceding column. Each *S* served for 14 or more trials by repeating this 13 by 13 square.

---

[5] E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951, 540-545.

[6] R. E. Bowman, Serial and temporal position-effects in rote learning, M. S. thesis, University of Wisconsin Library, 1956, 53; R. E. Bowman and W. R. Thurlow, Serial and presentation position effects in rote learning, *Amer. Psychologist,* 11, 1956, 450.

Serial positions were established by instructing *Ss* to conceive of the syllables as if serially ordered in a vertical column on a sheet of paper, with the 'first' syllable coming at the top. Each *S* in a group was given a different one of the 13 syllables as his 'first' syllable of the list; since order of syllables was the same for all *Ss*, this procedure counterbalanced syllables at each serial position. The *Ss* were told that each trial would start with a different syllable of the list, proceeding in order from that starting point on through the list, and that when the last syllable was reached, the first syllable would follow next.

If, upon post-experimental questioning, *S* reported as 'first' a syllable different from the syllable he had been instructed was first, his data were discarded and another *S* was substituted for him. Eleven such instances occurred in the course of the experiments with Groups 1, 2, and 3. Data were also excluded, and additional *Ss* tested, for 17 *Ss* who failed to achieve an *a priori* learning criterion of 100 correct responses in 26 trials in the four experimental groups. A graphical examination of the data excluded indicated that they were similar to those reported below, despite low frequencies of response and the lack of counterbalancing of syllables, serial, and within-trial positions.

The memory-drum used was timed to present each syllable for 2 sec., followed by a 2-sec. interval during which the shutter was closed and the drum positioned the next syllable. The consecutive exposure of all 13 syllables constituted a trial. During the inter-trial interval of 12 sec., the shutter was closed, and the *Ss* were required to count digits aloud to minimize rehearsal of syllables. The *Ss* spelled the syllables.

Details specific to each Experiment and Group are given below.

Experiment I. List A was used. The *Ss* were not required to spell the syllable when exposed if they had just correctly anticipated it. The 'first' syllable was identified in the instructions as that syllable which would be seen with a blue background around it. This 'first' syllable was presented first on Trial 1 for all *Ss*. Two groups were used. For Group 1, the 'first' syllable provided the only anchor. For Group 2, there were in addition orange backgrounds at the fifth and ninth serial positions, to induce the *Ss* to learn the syllables in clusters.

Experiment II. List B was used. The *Ss*, furthermore, spelled each syllable upon its exposure, regardless of the correctness of their anticipation, which held relatively constant the time available for anticipating the subsequent syllable. No colored anchors were used: instead, each *S* was shown his 'first' syllable for 2 sec. on the memory-drum during his instruction. Two groups were used. For Group 3, Trial 1, began with a different serial position for each *S*. Thus, for each trial, serial positions were counterbalanced at each within-trial position. For Group 4, the serially first syllable occurred first on every trial for each *S*, thus providing comparative data on the effect of serial position as obtained in the classical manner.

Experiment III. Only one group (Group 5) was used. The procedure was identical to that for Group 3, except that the first five syllables of each *S*'s list were printed in one color, the next four in a second color, and the last four in a third color. The *Ss* were instructed that the syllables formed three short lists, each printed in a different color. No mention of a 'first' syllable was made, and thus no data on serial position are available from this group.

Results: (1) Effect of serial-position. Total correct responses for each group at each of the serial positions were determined. Data for Trails 2 through 14 were used to obtain a block of trials over which counterbalancing of serial positions at each within-trial position was complete. These data were analyzed for statistical significance using Grant's extension of the Alexander trend-test.[7]

The analysis of the two serial-position curves from Experiment I (Table I) demonstrates the over-all effect, but no statistically significant difference between the two curves, except at the 5% level for the residual of all trends above the cubic. Graphic inspection of the data (Fig. 1) also indicates the similarity of these two curves. It thus appears that the attempt in Group 2 to structure the fifth and ninth syllables as anchors failed, and the *Ss* learned the list in the same way as did the *Ss* of Group 1.

TABLE I

SUMMARY OF *F*-VALUES FROM TREND ANALYSIS ON SEPARATE EXPERIMENTS

Source of variance	Effect of serial position			Effect within trial position		
	'First' syllable First		'First' syllable last			
	Experiment I Grp. 1 *vs.* Grp. 2	Experiment II Grp. 3 *vs.* Grp. 4	Experiment II Grp. 3 *vs.* Grp. 4	Experiment I Grp. 1 *vs.* Grp. 2	Experiment II Grp. 3	Experiment III Grp. 5
Overall trend	17.93‡	11.64‡	12.72‡	8.81‡	3.04†	6.03‡
linear	88.53‡	31.01‡	15.73‡	2.42	4.33	6.48*
quadratic	19.44‡	44.62‡	84.56‡	18.14‡	2.77	9.35†
cubic	3.95	0.22	0.00	65.57‡	11.38†	23.24‡
residual	10.19‡	1.31	1.09	3.07†	1.69	4.41‡
Between group means	0.12	45.66‡	45.66‡	0.17	—	—
Between group trends	1.45	2.07*	1.00	0.72	—	—
linear	0.74	0.09	0.69	0.12	—	—
quadratic	0.40	7.22*	1.89	3.50	—	—
Cubic	0.03	0.00	0.33	0.13	—	—
residual	2.31*	1.93	1.08	0.38	—	—
Between individual Means	7.55‡	2.72‡	2.72‡	18.89‡	11.50‡	22.85‡

* $p=0.05$. † $p=0.01$. ‡ $p=0.001$.

The serial position curves from Experiment II offer a quantitative comparison of the effects obtained by the experimental vs. the classical method of presenting the serial list. As seen in Fig. 1, the curve for the classical

---

[7] D.A.Grant, Analysis-of-variance tests in the analysis and comparison of curves, *Psychol. Bull.*, 53, 1956, 141-154.

Control Group (Group 4) presents the typical bowed effect of serial position. The curve for Group 3, and also for Groups 1 and 2 in Experiment I, present one major difference from this: the syllable most frequently anticipated correctly was the second in the list rather than the first. This feature probably accounts for the significant difference in the quadratic component of trend between the classical control and experimental curves of Experiment II (Table I). The same difference is also noted in the over-all analysis of Groups 1, 2, 3, and 4 (Table II).

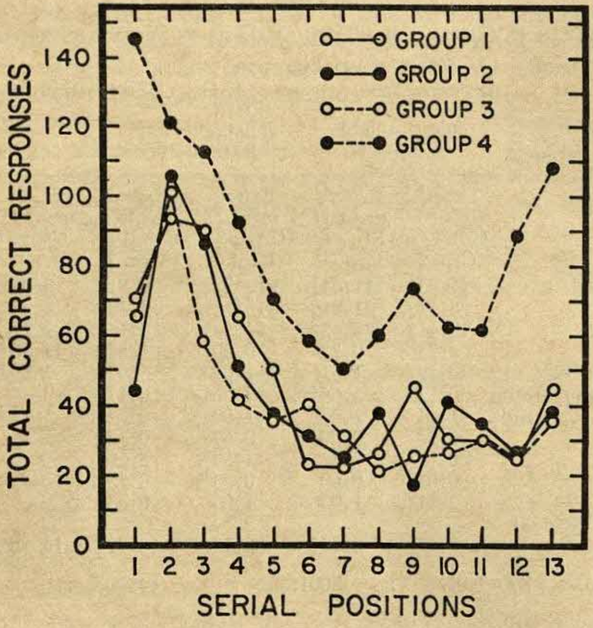


FIG. 1. EFFECT OF SERIAL POSITION UNDER THE CLASSICAL AND EXPERIMENTAL CONDITIONS

The following explanation for the above difference is offered. In the Control Group, the inter-trial interval of 12 sec. was always succeeded by the serially first syllable, and thus served as the stimulus for the first syllable. As a stimulus, the inter-trial interval may be regarded as a heterogeneous item in an otherwise homogeneous series of syllable-stimuli. Conversely for the Experimental Groups, the first syllable of the list occurred on 12 of the 13 trials immediately following the last syllable rather than following the inter-trial interval. The *S*s in these groups therefore had to learn the last syllable, a homogeneous item, as the stimulus to anticipate the first syllable. For homogeneous items, this 'backward learning' progressed more slowly than the 'forward learning' of anticipating the second syllable following the first, in line with the results demonstrated by Ribback and Underwood.[8]

If this explanation is correct, then it would be reasonable to regard the 'first' syllable as 'last' for the Experimental Groups, thereby eliminating the effect of 'backward learning' from the beginning of the serial list. If this is done, then it can be visualized from the data of Fig. 1 that the only difference of note remaining between the Control and Experimental Groups is that the former demonstrates more correct responses at all points. In agreement with this observation, a trend-analysis, performed on the data treating the instructed 'first' syllable as 'last,' indicates that there is no significant difference between the Experimental and Control groups of Experiment II, in either the linear, quadratic, or cubic trends, or the residual (Table I). The difference between the means is however, highly significant. Thus, the experimental procedure of counterbalancing serial positions at within-trial positions has yielded a serial-position curve the shape of which is quantitatively similar to the curve obtained by the classical method of presentation. As might be expected from an examination of the data of Fig. 1, the same results are obtained in the over-all

TABLE II

Summary of *F*-Values from Trend Analyses on all Experiments Combined

Source of variance	Effect of serial-position (Groups, 1, 2, 3, and 4)		Effect of within-trial position (groups 1, 2, 3, and 5)
	'First' syllable first	'First' syllable last	
Overall trend	24.42‡	26.83‡	15.67‡
linear	95.68‡	34.73‡	11.22†
quadratic	63.31‡	205.63‡	29.85‡
cubic	2.78	1.35	92.07‡
residual	6.82‡	1.86	6.75‡
Between group means	17.63‡	17.63‡	0.25
Between group trends	1.79*	0.99	1.02
linear	0.25	0.38	0.44
quadratic	4.77†	1.34	2.26
cubic	0.33	1.33	2.06
residual	1.84*	1.10	0.78
Between individual trends	3.53‡	3.53‡	18.20‡

* $p=0.05$. † $p=0.01$. ‡ $p=0.001$.

---

[8] Alan Ribback and B. J. Underwood, An empirical explanation of the skewness of the bowed serial position curve, *J. exp. Psychol.*, 40, 1950, 329-335.

analysis of Groups 1, 2, 3, and 4 (Table II), treating the 'first' syllable as 'last.' It may thus be stated that the four procedural differences between Experiments I and II (change of list, spelling of correctly anticipated syllables upon exposure, method of emphasis by anchor-syllables, and completeness of counterbalancing) have together had no effect on the curves obtained for serial position. It cannot be said whether this is the result of a lack of effect of each of these differences singly, or whether two or more of these procedural changes have exerted mutually cancelling effects, which would seem to be unlikely, but must be considered as a possibility. In either event, it is clear that the same experimental effect of serial-position has been obtained under two defined experimental procedures, and this is the point of major concern here.

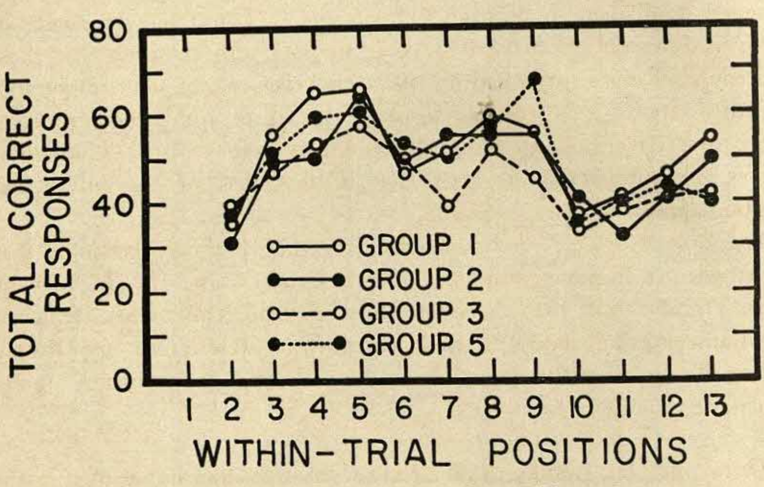


FIG. 2. THE POSITIONAL EFFECT OBTAINED WITHIN TRIALS UNDER THE EXPERIMENTAL POSITIONS

It may also be noted in Table I and II that the over-all trend for the curves of serial-position, for all of the analyses performed, exhibits highly significant linear and quadratic components, but not cubic. These first two components would be expected to be significant from the well established simple concavity of the curve of serial-position, and from the general finding that correct responses are lower for the end than for the beginning of the list.

(2) Effect of within-trial position. Fig. 2 displays the total correct responses summed over Trials 2 through 14 for each of the experimental

groups as a function of within-trial positions. For each of these groups over this block of trials, each serial position is represented equally often at each within-trial position, and thus the shape of the curves should be independent of the serial-position effect.

The four curves for within-trial positions have a complex but fairly similar shape, appearing to pass successively through a maximum, a minimum, a maximum and a minimum. Although the separate analyses of Experiments I, II, and II do not all find the same significant components of trend, neither the two groups of Experiment I analyzed together (Table I) nor all four groups analyzed simultaneously (Table II) differ significantly in any component of trend. Since several procedures were changed simultaneously between Experiments I, II, and III, the results of these analyses do not make it possible to state with certainty that these procedural changes singly had no effect. However, and of most importance for this report, it is possible to state that three similar experimental procedures have been defined, all of which gave statistically identical curves for within-trial position. That is, the effect of the within-trial position appears to be reproducible.

For Groups 1, 2, 3, and 5 analyzed together, it should be noted that all trends are highly significant (over-all trends, Table II). In view of Grant's observation that all components of trend tend to be significant for transcendental functions, this generality of significance may simply reflect the fact that the shape of the over-all curve approximates a sine-function.[9]

Discussion. The finding of an effect of serial position under the present conditions, in which each syllable was presented at a different position from trial to trial, is difficult to reconcile with most existing theories of that effect. For example, consider the effect predicted for Group 3 by the mathematico-deductive theory of rote-learning. This theory states that inhibitory potentials (which account for the effect of serial position) develop and accumulate as a function of the position of the syllables within the trial over successive trials. In Group 3, the syllable occupying position *n* on Trial 1 occupied the other 12 positions on the next 12 trials in the same order for all *S*s. Therefore, the accumulation of inhibitory potentials to that syllable on any trial was the same for all *S*s. Since any syllable represented a different serial position for each *S*, then on any trial the inhibitory potentials for each serial position for the group were the sum of the same 13 potentials, *i.e.* were identical. In other words, the theory predicts no serial position effect.

---

[9] Grant, *op. cit.* 144.

Similarly, theories based on proactive and retroactive inhibition fail to account for the present effect of serial position, since all syllables of the list should accumulate equally the effects of both forms of inhibition, occurring as they do at all possible degrees of remoteness from the beginning syllable of the trial. In like manner, theories postulating competing responses as a function of remote forward and backward associations would seem to suffer the same defect of having to predict equal amounts of such associations, and therefore response-competition for all serial positions in the present case.

One theoretical explanation of the effect of serial position that is not at variance with the present results is that of association with position,[10] if such association is understood to be with list-positions rather than trial-positions. This hypothesis has never, however, been developed, and is not sufficient by itself. Another possible explanation of the present results is that they are examples of the Von Restorff phenomenon: the syllable emphasized as first by the instructions served as an 'anchor' about which the learning progressed. This progression could then be thought of in the form of forward and backward learning of the type demonstrated by Ribback and Underwood. These possibilities are without proof, but suggest directions for future experimentation.

The interpretation of the within-trial effect of position is not completely clear, although a ready explanation exists for the shape of the curve from Positions 2 through 5. The *S*s did not know which syllable of the serial list would begin each trial, and after seeing the starting syllable of the trial, had then to anticipate the serially next syllable with only the immediately preceding syllable as the stimulus. Each additional presentation of a syllable, up to the fifth, increased the efficiency of anticipation, probably by building up the multiple-stimulus complex to which the subsequent syllable was conditioned. The sine-like oscillation subsequent to Position 5 is not, however, so readily explained. It is not even clear whether this oscillation is independent of the effect of serial position. In one experiment, in which syllables were fixed at spatial positions in a circular array, and presented in non-serial order, the effect of position within a trial had such a sine-like appearance.[11] On the other hand, subsequent experiments with the same list of syllables on a memory-drum indicated that the curve of position within trial was altered somewhat by changing the systematic order with which starting syllables were chosen for each trial, and thus may have been influenced by the effect of serial position concurrently present. In line with this, a prediction developed from the *Mathematico-Deductive Theory of Rote Learning,* using constants derived in that book from classical data on serial position, indicates such an oscillating effect of position within trials from Positions 5 through 13 (although the prediction fails for the first portion of the curve by indicating that correct responses should begin high, and then decrease). Clarification of this matter will require further experimentation.

---

[10] McGeoch and Irion, *op. cit.*, 134.

[11] Bowman, *op. cit.*, 59.

SUMMARY

Most theories of the effect of serial position depend upon the congruence of the serial order of the list with the serial order of presentation on the trials. This experiment investigated whether such congruence is necessary for obtaining the effect of serial position. Lists of 13 syllables in fixed, circular sequence on a memory-drum were presented by the anticipation method to *S*s, each of whom was instructed that a different syllable of the list was 'first.' This 'first' syllable was then presented once at each of the 13 within-trial positions over the subsequent 13 trials, always in its proper place in the circular sequence of syllables.

Under these conditions, an effect of serial position was obtained in which the 'second' serial position had the highest frequency of correct responses. If this is interpreted in terms of 'forward' and 'backward' learning, the 'first' syllable must be regarded as the 'last' to eliminate 'backward' learning from the beginning of the serial list. There was no significant difference between this experimental effect and the classical effect obtained with a control group. These results suggest that list-position, not trial-position, is the important determinant of the effect of serial position, and thus are counter to the explanations offered by the major theories.

In addition, a within-trial position-effect was observed, in which the increase in responses up to the fifth trial-position suggested that more than one syllable was utilized as the stimulus for the subsequent anticipation. A subsequent sine-shaped oscillation in frequency of correct responses frequency was noted.

SCALE-SYMMETRY AND THE SEMANTIC DIFFERENTIAL

By EDWARD J. MALMSTROM and GILBERT M. FRENCH,
University of California, Berkeley

In discussing the general properties of rating scales used in semantic differentials, Osgood, Suci, and Tannenbaum make two basic assumptions, which they discuss under the heading of 'linearity.'[1] The assumptions are that the polar adjectives of each scale are represented by points in semantic space that lie at the same distances but in opposite directions from the origin. If both assumed properties are present, concepts of exactly opposite meaning should show mirror-image profiles when rated on the same semantic differential scales. In other words, the rating scales are symmetrical with equal distances between successive numbers. If these assumptions are correct, it should be possible to predict the judgments on the usual bipolar scales by adding judgments obtained with two opposing unipolar scales, each having one end at the origin. This approach to the problem was taken in an unpublished study by Taylor and Kumata, which has been reported briefly by Osgood, Suci, and Tannenbaum.[2] Each of four concepts was judged both in the customary manner on 10 bipolar 7-point scales and unipolarly on 4-point scales identified by only one of a pair of adjectives. If for a given concept the *Ss*' mean judgments were 2 and 1, respectively, on unipolar scales *good* and *bad*, the algebraic sum of the signed values (+2 and −1) would yield the expected value of +1 for the associated mean bipolar judgment. Summed unipolar results were not significantly different from results obtained directly from corresponding bipolar scales in 90% of the possible experimental comparisons.

In the present investigation, a different approach to the assessment of symmetry was sought. The method involved precise control over the information available to *Ss* about the concepts to be judged. Information was given in terms of ratings on two 5-point scales of a semantic differential and was varied in scale-value and direction, positive or negative. For the concept thus described, *Ss* made ratings on the remaining scales. It was expected that increased value on the information-scales would be

---

* Received for publication August 25, 1961.

[1] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, *The Measurement of Meaning*, 1957, 152-153.

[2] Osgood, Suci, and Tannenbaum, *op. cit.*, 53-57.

reflected directly in a linear increase in mean judgments. Since the information-scales always had their primary loadings on Osgood's Evaluative Factor, it was predicted that the polarity of judgments would correspond consistently with the polarity of information only for other evaluative scales. Another prediction was that with discrepant or conflictful information, ratings would be lower, on the average, than with non-discrepant or compatible information. For instance, if a concept were described as both *good* and *dirty*, judgments would be more neutral than if it were described as both *good* and *clean*.

METHOD

Subjects. Eighty-one undergraduate students were drawn from several courses in psychology at the University of California, Berkeley. None of the 21 men and 60 women who participated had had previous experience with semantic differentials.

Semantic differential. Each S was required to rate 9 concepts on 19 of the 21 scales listed in Table I. The concepts were identified as 'persons,' about whom the information was presented by means of ratings on the *good-bad* scale and one of the other scales italicized in the table; the second information-scale was different for every concept to be judged by the S. The information-scales were selected because of their relatively high loadings on the Evaluative Factor and because of their particular relevance to judgment of persons.

TABLE I

ADJECTIVAL SCALES USED IN THE SEMANTIC DIFFERENTIAL

The factor on which a scale has its major loading, according to the analysis of Osgood, Suci, and Tannenbaum, is indicated by the letter in parentheses: E for Evaluative, P for Potency, and A for Oriented Activity.

(1) *good-bad* (E)
(2) *wise-foolish* (E)
(3) *clean-dirty* (E)
(4) *sociable-unsociable* (E)
(5) *successful-unsuccessful* (E)
(6) *positive-negative* (E)
(7) *beautiful-ugly* (E)
(8) *grateful-ungrateful* (E)
(9) *kind-cruel* (E)
(10) *reputable-disreputable* (E)
(11) honest-dishonest (E)
(12) graceful-awkward (E)
(13) masculine-feminine (P)
(14) deep-shallow (P)
(15) serious-humorous (P)
(16) strong-weak (P)
(17) large-small (P)
(18) tenacious-yielding (P)
(19) active-passive (A)
(20) excitable-calm (A)
(21) fast-slow (A)

Five-position scales were used. The extreme positions (+2 and −2) were labelled as "very well describes the person." The next more central positions (+1 and −1) were labelled as "fairly well describes the person." The most central position (0) was labelled as "neutral, both sides equally descriptive, or completely irrelevant." For each of his nine series of judgments, an S was given a different combination of ratings on *good-bad* and the other information-scale. The various combinations were: both +2, both −2, both +1, both −1, both 0, +2 and −2, +1 and −1, −2 and +2, and −1 and +1, for the respective scales. The first five combinations yielded compatible information and the last four discrepant information.

Over the course of the entire experiment, each of the sets of two information-scales appeared nine times with each combination of scale-positions. From S to S, the order of presentation of these conditions was varied at random. Also, for every 'person' rated by an *S*, the 21 scales of the semantic differential were presented in a different random arrangement with polarity of scales randomly reversed. These procedures were designed to minimize possible habitual modes of responding and to draw the *Ss'* attention to the individual judgments.

Before booklets with the sematic differentials were distributed to the *Ss*, instructions were given both orally by *E* and in writing on separate sheets. The instructions were patterned closely after those described by Osgood, Suci, and Tannenbaum,[3] except that there was an emphasis on the making of judgments about 'persons' and on the special manner in which the information was to be presented.

TABLE II

RESULTS BASED ON ABSOLUTE VALUES OF JUDGMENTS FOR DIFFERENT COMBINATIONS OF VALUES ON INFORMATION-SCALES

Values on information-scales*	Over-all mean ratings	t, (19 df)	r
Both +2	1.13		
Both −2	1.14	0.11	+.74
Both +1	0.90		
Both −1	0.84	0.72	+.76
+2 & −2	0.91		
−2 & +2	1.00	1.47	+.86
+1 & −1	0.78		
−1 & +1	0.82	0.61	+.79

* For those conditions in which information of opposite polarity was given on the two scales, the sign of the information on the *good-bad* scale is given first.

The booklets contained a semantic differential on each of nine pages. The consecutive pages of a booklet were headed "Person No. 1," "Person No. 2," and so on. Scale-positions were indicated by Xs. To contrast with the black printing of the pages of the booklet, both information-scales on a page were marked with a red penciled circle around the appropriate position. Because of the differing random arrangement of the scales within each semantic differential, the information-scales appeared at a different position on each page. At the bottom of a page appeared the admonition, "Make sure you have checked every scale."

RESULTS AND DISCUSSION

For each combination of values on the information-scales, means of the absolute values of *Ss'* ratings on every judgment-scale were computed. For those reciprocal conditions in which the information was of opposite polarity, product-moment correlations were obtained between the means for the 20 judgment-scales. (Since *good-bad* was always an information-scale, no judgments were obtained for it.) The correlations, shown in the final column of Table II, are all positive and significant at beyond the 0.1% level. Comparisons of the grand means of ratings made under the reciprocal conditions are also given in Table II where it can be seen readily that none of the differences is statistically significant. Pooled results for each pair of reciprocal conditions, along with the mean rating obtained for information of a scale-value of 0, are shown in Fig. 1. Separate curves have been drawn for discrepant and non-discrepant information. An *F*-test of linearity reveals that the two functions are not significantly quadratic in form.[4] Under the present experimental conditions, then, the scale-value of judgments increases in a simple linear manner as the scale-value of information increases. The hypothesis of lower scale-value of judgments with discrepant information is supported by the data. The difference between mean ratings for discrepant and non-discrepant information of scale-value 2 is significant at better than the 0.1% level ($t = 2.96$, 80 *df*); and the significance of the difference for scale-value 1 lies between the 10% and the 5% levels ($t = 1.47$, 80 *df*).

---

[3] Osgood, Suci, and Tannenbaum, *op. cit.*, 82-84.

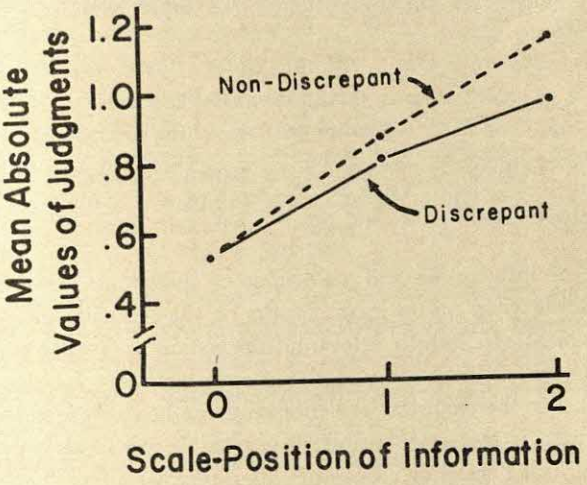


FIG. 1. MEANS OF ABSOLUTE VALUES OF JUDGMENTS FOR DISCREPANT AND NON-DISCREPANT INFORMATION OF DIFFERENT ABSOLUTE SCALE-VALUE

While polarity of judgment corresponded closely with polarity of information for evaluative rating scales, judgments on those adjectival scales highly loaded on the factors of Potency or Activity were often neutral or of opposite polarity. Table III illustrates this outcome, showing mean signed values of judgments on evaluative and non-evaluative scales for non-discrepant information of varying polarity and value. The range of mean values of judgments obtained on evaluative scales is much larger than that for the judgments on non-evaluative scales. Also, the results for evaluative scales show a good agreement of polarity with polarity of information. If there is any trend for the non-evaluative scales, the mean of judgments is somewhat greater with greater values of information, no matter what the polarity. This is in keeping with general results described previously.

---

[4] J. P. Guilford, *Fundamental Statistics in Psychology and Education,* 3rd ed., 1956, 294.

Taken together, the findings argue for a direct relationship between scale-value of information and scale-value of judgments. Concepts of opposite meaning yield mirror-image results on the semantic differential for those judgments for which relevant information is available. The data support the two assumptions underlying scale-symmetry for the particular scales used. But the empirical evidence must not be overgeneralized. What is shown here is that scale-symmetry is possible with pairs of polar adjectives. For other scales or groups of scales, symmetry must be determined independently.

TABLE III

MEAN SIGNED JUDGMENTS BY TYPE OF SCALE

Type of scale	Polarity and value of the information				
	Both −2	Both −1	Both 0	Both +1	Both +2
Evaluative	−1.12	−0.75	−0.08	+0.91	+1.31
Non-evaluative	+0.14	+0.07	−0.22	+0.05	+0.27

The finding that mirror-imaging of results is obtained consistently only when the scale of the required judgment has its primary loading on the same factor as the scales of information is, perhaps, worthy of more than routine interest because it suggests an alternative interpretation of similar results obtained by Asch.[5] His *S*s were presented the following set of adjectives descriptive of a 'person': *intelligent, skillful, industrious, determined, practical, cautious.* One group of *S*s was given in addition the adjective *warm* and another group *cold.* The *S*s rated the 'person' by choosing one adjective from each of 18 pairs on a checklist. Some, but not all, of the judgments were reversed in direction when *cold* was given instead of *warm.* The tendency toward reciprocal profiles was much weaker in another experiment in which *polite* and *blunt* were substituted for *warm* and *cold,* an outcome Asch interpreted as indicating that the former adjectives are less 'central' (important) for formation of impression of persons than are the latter.

---

[5] S. E. Asch, Forming impressions of personality, *J. abnorm. soc. Psychol.,* 41, 1946, 258-290.

Wishner reëvaluated Asch's data and concluded that a stimulus-adjective will have strong effects if the items to be judged are correlated with it and weak effects if they are uncorrelated.[6] It is suggested here, in addition, that the differences in factor of major loading and in magnitude of loading account for Asch's results. On the basis of Osgood's published data,[7] it is possible to identify in the list used by Asch those pairs of adjectives with loadings in the Evaluative Factor. When *cold* took the place of *warm* in the stimulus-list, 8 of the 11 evaluative adjectival pairs produced opposite results, while 6 of the 7 non-evaluative pairs failed to yield such reversals. One thus has a good basis for assuming that *cold* and *warm* are 'evaluative' in the context in which Asch used them. No clear trend appears in the findings with *polite* and *blunt,* suggesting that these adjectives are non-evaluative or only slightly evaluative in character.

SUMMARY

Eighty-one *S*s judged 9 concepts on 19 of the 21 scales of a semantic differential. The only information available about each concept was given on the two remaining scales and was systematically varied in scale-value and polarity. Another dimension of variation was discrepant versus compatible information.

For information of the same scale-value but opposite polarity, there was a high, positive correlation between means of absolute values of judgments on the various scales; and the grand means were not significantly different. The data also showed that the scale-value of judgments varies as a simple linear function of the absolute scale-value of information. Since all the information was given in terms of scales with primary loadings on the Evaluation Factor, polarity of judgments was consistent with polarity of information only for other evaluative scales. Compatible information yielded judgments of higher scale-value than did discrepant information.

---

[6] Julius Wishner, Reanalysis of "Impressions of Personality," *Psychol. Rev.,* 67, 1960, 96-112.

[7] Osgood, Suci, and Tannenbaum, *op. cit.,* 31-75.

SIZE- AND DISTANCE-ESTIMATES OF A SINGLE STIMULUS UNDER DIFFERENT VIEWING CONDITIONS

By RAY OVER, University of Sydney

It is proposed to examine the effect of variation in stimulus-size and stimulus-distance on both judgments and estimates of size and distance. These two measures can be characterized in terms of different arrangements of stimuli and different instructions. *Judgments* are obtained when *S* is presented with *two stimuli* and instructed to relate their sizes and distances, *estimates* when *S* is prepared with *a single stimulus* and asked to specify its size and distance.

Under reduced viewing conditions, two stimuli are judged equal in both size and distance when they subtend the same visual angle.[1] Under unrestricted viewing conditions, both size-constancy and distance-constancy are found;[2] two stimuli of the same physical size are judged equal in size irrespective of differences in their distances from *S*, and two stimuli at the same physical distance from *S* are judged equal in distance despite differences in their sizes. There have been complaints that different methods of judgment yield limited information about *S*'s perception and that judgments of size- and distance-equality do not show whether *S* perceives the stimuli as small, large, near, far, or how small, how large, etc. Four procedures have been used to obtain such information—modified methods of judgment, ratio methods of judgment, single-stimulus rating with an arbitrary scale, and single-stimulus estimation with an absolute zero and equal-unit scale.[3] The present concern will be with this last method in which *S* is asked, "What size is this 'single' stimulus and how far is it away from you?"

Some relationships between size- and distance-estimates and the physi-

---

* Received for publication September 20, 1961.

[1] Ray Over, Size and distance judgments under reduction conditions, *Austral J. Psychol.*, 12, 1960, 162-168.

[2] A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192; Over, Distance-constancy, this JOURNAL, 74, 1961, 309-310.

[3] W. H. Ittelson, Size as a cue to distance, this JOURNAL, 64, 1951, 54-67, 188-202; H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-426; V. V. Lloyd, The interaction of stimulus area and intensity as cues in the perception of distance, *J. gen. Psychol.*, 49, 1953, 167-183; Egon Brunswik, Distal focussing of perception: Size-constancy in a representative sample of situations, *Psychol. Monog.*, 56, 1944, (No. 254), 1-49.

cal size and distance of a single stimulus can be predicted from data based on the method of judgment. Under unrestricted viewing conditions, the estimated size of a single stimulus should correspond to its physical size (*i.e. S* should report "three inches" when asked to estimate the size of a 3-in. stimulus) and estimated distance should correspond to physical distance. Under reduced viewing conditions, size- and distance-estimates should be a function of *both* physical size and distance.

Boring's and Gibson's discussion of this latter possibility was to some extent at cross-purposes in that many of Boring's remarks applied to the judgment-situation and many of Gibson's to the estimate-situation.[4] Gibson stated that, "The effect of stimulus-reduction on object-perception is to substitute for the normal perceptual process of size-determination an attitudinal process. . . . When distance is not determined by stimulation, it may be determined by some *presumed* distance on an imaginal plane. . . . When a presumed distance arises in experience a presumed size accompanies it."[5] Gibson has used the term 'indeterminate' to describe this process, and this has led Boring to reply, "Gibson is in error in thinking that stimulus-controlled distance, under reduction, becomes indeterminate . . . (that) seen size (gets) more and more independent of retinal image size with increasing elimination of clues to distance."[6] It is possible that by "indeterminate" Gibson has meant simply that an estimated size will not necessarily be the same as a physical size nor an estimated distance the same as a physical distance. The estimates will be determinate, however, in the sense that when a particular distance-estimate is given by *S* the size-estimate will be determined by the visual angle of the stimulus. The two estimates will be made with one degree of freedom; they will be indeterminate in that neither will necessarily correspond to the respective physical values of the stimulus but they will be determinate in that their ratio will equal the ratio of the two physical values.

The purpose of the study reported here is to examine the effect of variation in the physical size and distance of a single stimulus-object on estimates of size and distance obtained under both reduced and unrestricted viewing conditions.

METHOD

Apparatus and procedure. Four *S*s, tested individually, each made 180 estimates of the size and distance of a single stimulus, for which the combinations of physical size and distance were varied between trials. The first 90 pairs of estimates were obtained under reduced viewing conditions, the second under unrestricted viewing conditions.

Reduced viewing conditions. The testing situation was similar to that in which judgments of equality of size and distance were given when two stimuli subtended the same visual angle.[7] The restrictions on viewing included the introduction of *S* to a darkened room, monocular viewing, and the restriction of head-movements. The stimulus was a diamond-shaped square of light cast on a milk-glass screen, the luminous intensity of which was constant at 0.0457 ft.-L. Three sizes—2-, 3-, 4-in. diagonals—and three distances—10, 20, 30 ft.—provided nine combinations of size and distance.

---

[4] E. G. Boring, The Gibsonian visual field, *Psychol. Rev.,* 59, 1952, 246-247; J. J. Gibson, The visual field and the visual world: A reply to Professor Boring, *Psychol. Rev.,* 59, 1952, 149-152.

[5] Gibson, *op. cit.,* 151.

[6] Boring, *op. cit.,* 247.

Before *S* entered the testing situation, he was instructed:

Instructions. When you see a diamond-shaped light source you are to estimate in inches the reading you would obtain if you were to walk out to this object and place a tape-measure across its horizontal diagonal. You are also to estimate in feet the physical distance of the object from you.

TABLE I

MEAN ESTIMATES OF SIZE AND DISTANCE IN INCHES AND FEET UNDER DIFFERENT VIEWING CONDITIONS

		Conditions					
		Reduced viewing			Unrestricted viewing		
		stimulus-distance			stimulus-distance		
Estimated	Stimulus size (in.)	10 ft.	20 ft.	30 ft.	10 ft.	20 ft.	30 ft.
Size	2	3.05 in.	2.61 in.	2.30 in.	1.61 in.	1.57 in.	1.63 in.
	3	3.73 in.	3.05 in.	2.58 in.	2.51 in.	2.51 in.	2.60 in.
	4	3.88 in.	3.16 in.	2.96 in.	3.53 in.	3.58 in.	3.60 in.
Distance	2	14.30 ft.	19.88 ft.	23.05 ft.	9.46 ft.	18.78 ft.	27.83 ft.
	3	11.68 ft.	16.83 ft.	20.25 ft.	9.45 ft.	18.54 ft.	27.80 ft.
	4	9.25 ft.	15.53 ft.	17.60 ft.	9.46 ft.	18.73 ft.	28.20 ft.

Each *S* made an estimate of size and distance on each of the 10 presentations of the 9 combinations of size and distance of the stimulus, the order of trials within the 10 sets being varied at random.

Unrestricted viewing conditions. Testing was carried out in the same room which was now well illuminated. Viewing was binocular, and *S* was allowed to make head-movements. The stimulus was a white diamond-shaped square of cardboard fitted to a tripod-stand at eye-level. The combinations of size and distance of the stimulus and the order of presentations were the same for the two viewing conditions.

RESULTS

The mean estimates of size and distance given under the different viewing conditions for the nine combinations of size and distance of the stimulus-object are given in Table I. Because of differences in the variances of the subgroups it was not possible to consider the effect of the single and joint variation of the independent variables on the response-measures by analyses of variance. Some qualitative comparisons can, however, be made.

---

[7] Over, The effect of instructions on size-judgments under reduction-conditions, this JOURNAL, 73, 1960, 599-602.

Under unrestricted viewing conditions, increases in the size of a stimulus at a constant distance from *S* led to increases in size-estimates but to no change in distance-estimates. Increases in the distance of a stimulus of constant size led to increases in distance-estimates but to no change in size-estimates. Under reduced viewing conditions, both an increase in the size and a decrease in distance led to changes in estimates of both size and distance. When the results on which the means in Table I are based are considered, the mean of the various values of estimated size/stimulus-size is 1.10 (*SD* = 0.55) under reduced viewing conditions, and 0.91 (*SD* = 0.20) under unrestricted viewing conditions. The mean of the various values of estimated distance/stimulus-distance is 0.84 (*SD* = 0.48) under reduced viewing conditions, and 0.98 (*SD* = 0.18) under unrestricted viewing conditions. The critical difference in results between the two viewing conditions is in the *SD*s. Under unrestricted viewing conditions, although there was a trend toward the underestimation of size and distance, there were few large errors in estimation. Under reduced viewing conditions, however, more than half of the values fell outside the range 0.75-1.25; *i.e.* the stimulus was often estimated to be much smaller, larger, closer, or farther away, than it really was.

If *S* underestimated (or overestimated) size by a given degree, it can be asked if he also underestimated (or overestimated) distance by the same degree. When the results of which the values in Table I are means are considered, there is a correlation of +0.35 ($\sigma_{s.d} = 0.07$, $\sigma_{d.s} = 0.06$) under unrestricted viewing conditions, and of +0.82 ($\sigma_{s.d} = 0.46$, $\sigma_{d.s} = 0.41$) under reduced viewing conditions, between the values of estimated size/stimulus-size and the values of estimated distance/stimulus-distance. The two correlations cannot be compared because of different standard errors. A more meaningful comparison of the relationship between estimated size and estimated distance can be found in Table II where the frequency-distribution of the ratios of estimated size/stimulus-size to estimated distance/stimulus-distance are set out for the two viewing conditions. Errors of size-estimation of a given degree are much more related to errors of distance-estimation of the same degree under unrestricted viewing conditions than under reduced viewing conditions; *i.e.* estimate-ratios, as well as size-estimates and distance-estimates, are much more determinate under unrestricted viewing conditions than under reduced viewing conditions.

DISCUSSION

When *S*s are required to make size- and distance-judgments of two stimuli under reduced viewing conditions, the relational information (two stimuli subtending visual angles which either are equal or unequal) on which judgments are based is provided in the situation. This is the only information *S* can use in making his judgments, and it is not surprising that size- and distance-judgments follow the same relationship. In the present experiment *S* was presented with a single stimulus under reduced viewing conditions and explicitly instructed to estimate its physical size and distance. There was no explicit feature of the situation to which the single stimulus could be related. When questioned after the experiment, *S*s reported difficulty in making estimates. Although they had never been in the testing room nor seen the stimulus-objects before being tested under reduced viewing conditions, all *S*s stated that they formed 'hypotheses' about the possible length of the testing room and about the nature of the stimulus. They claimed to have made distance-estimates in terms of the former 'hypotheses,' estimates of size in terms of the latter. This would account for the particular ranges of sizes and distances given in the estimates, and it might be supposed that variation of estimates within these ranges was to some extent a function of the visual angle subtended by the stimulus. It is only when size- and distance-estimates are made with the explicit use of a single set of criteria that determinate estimate-ratios would be expected.

TABLE II

FREQUENCY-DISTRIBUTION OF RATIOS UNDER THE TWO VIEWING CONDITIONS

Viewing conditions	Ratios*								
	0.00–0.24	0.25–0.49	0.50–0.74	0.75–0.99	1.00–1.24	1.25–1.49	1.50–1.74	1.75–1.99	2.00+
reduced	0	1	11	7	6	0	2	1	8
unrestricted	0	0	0	9	16	8	2	0	1

* Values of (estimated size/stimulus-size)/(estimated distance/stimulus-distance).

Under reduced viewing conditions, it may be possible to get some *S*s to estimate that an object varying in visual angle is constant in size but varying in distance, and to get other *S*s to estimate that the same object is changing in size but is constant in distance. The particular function which estimates follow will depend on the 'hypotheses' *S*s use, and these presumably can be specified in terms of instructions and the past experience of *S*s. It is not known what procedures *E* would have to carry out such that size- and distance-estimates will be indeterminate and estimate-ratios determinate for all *S*s.

At first sight, the present results appear contrary to an hypothesis of size-distance invariance which states that discriminations of size and distance are related. It is a necessary requirement, however, for testing this hypothesis that the discriminations of size and distance are obtained under *equivalent* experimental conditions. In the present experiment, *S*s were instructed to use the one set of criteria to make both size- and distance-estimates, yet it is known that they did not. If *S*s *cannot* use the one set of criteria for both estimates under reduced viewing conditions, the more molecular form of the hypothesis of invariance cannot be tested. The more molar form of the hypothesis—which merely states that there is a relationship between discriminations of size and distance without stating that these must be obtained under equivalent conditions—is not confirmed by the present results.

SUMMARY

When *S*s estimate the size and distance of a single stimulus under unrestricted viewing conditions, size- and distance-estimates are determinate; *i.e.* they correspond to the physical size and distance, respectively, of the stimulus. Estimate-ratios (estimated size/estimated distance) are also determinate. Under reduced viewing conditions size- and distance-estimates are indeterminate, but an error in estimation of size in a certain direction is not necessarily accompanied by an error in estimation of distance in the same direction. It is suggested that this indeterminancy of estimate-ratios has arisen because size- and distance-estimates are made in terms of different relational scales.

PAIRED, ASSOCIATE LEARNING AS A FUNCTION OF RATE OF PRESENTATION AND PRIOR SERIAL LEARNING

By CHARLES C. ERICKSON, ROBERT D. INGRAM, and ROBERT K. YOUNG,
University of Texas

A series of studies was recently reported in which the effect of transferring from a serial (*Sr*) to a paired-associate (*PA*) task was investigated.[1] The results of each of these studies were in essential agreement: prior *Sr*-learning appears to have little influence upon the learning of a subsequent *PA*-task. Whether transfer is measured through the use of 'mixed' or 'unmixed' experimental designs, or through the use of positive or negative paradigms of transfer, items learned first in an *Sr*-list and then in a *PA*-list are learned no faster (or slower) than *PA*-items which have not been learned previously.

During the series of experiments mentioned above, the *S*s were asked why, in their opinion, *PA*-learning was not influenced by prior *Sr*-learning. Their responses were similar and may be summarized as follows: *S* had the *Sr*-list in mind while he was trying to learn the *PA*-list. When a *PA*-stimulus was presented, he would try to go through the *Sr*-list until he reached the *PA*-stimulus and then he would try to call out the next word in the *Sr*-list. The *PA*-list was presented so fast, however, that he would not call out the correct response in time.

According to the reports of the *S*s, the failure to demonstrate positive transfer under conditions when such would be expected was, at least partially, a function of the relatively rapid rate of presentation of the pairs during the *PA*-learning.

The *S*s' explanation of their method of *PA*-learning following *Sr*-training is tested in the present experiment. All previous research on the problem of transfer from *Sr*- to *PA*-learning has employed standard rates of presentation in both types of learning: a 2-sec. rate for *Sr*-learning and a 2:2-sec. rate for *PA*-learning. In the present experiment, rate of presentation during *Sr*-learning has been increased under the assumption that an increase in rate of *Sr*-presentation would enable the *S*s to recall the *Sr*-list with greater rapidity during the subsequent *PA*-learning. If the hypothesis as proposed by the *S*s has validity, then transfer from *Sr*- to *PA*-learning

---

* Received for publication January 2, 1962.

[1] R. K. Young, A comparison of two methods of learning serial associations, this JOURNAL, 72, 1959, 554-559; The stimulus in serial verbal learning, this JOURNAL, 74, 1961, 517-528; and Tests of three hypotheses concerning the stimulus in serial learning, *J. exp. Psychol.*, 63, 1962, 307-313.

should increase as rate of *PA*-presentation is decreased since this would allow the *S*s increasing amounts of time to recall the *Sr*-list for the needed *PA*-items. This prediction from the hypothesis is tested in the present experiment.

Method. The *S*s serving in the present experiment (48 in number) were taken from introductory psychology classes and were assigned randomly to one of three groups. Every *S* learned one of two *Sr*-lists of 14 items each (plus a cue item) and then all the *S*s learned a common *PA*-list composed of 14 pairs. Group I learned the *PA*-list at a rate of 1.5-1.5 sec. presentation, *i.e.* 1.5 sec. for the stimulus alone and then 1.5 sec. for the stimulus and response together, with a 3-sec. intetrial interval; Group II learned the *PA*-list at a 2:2-sec. rate of presentation with a 4-sec. intertrial interval; and Group III learned the *PA*-list at a 4:4 sec. rate of presentation with an 8-sec. intertrial interval. For all groups the *Sr*-list had a 1.5-sec. rate of presentation with a 4.5-sec. intertrial interval.

TABLE I

TRIALS TO LEARN *PA*-LIST
(N=16)

Group

I		II		III	
Mean	*SD*	Mean	*SD*	Mean	*SD*
18.13	10.33	11.31	4.97	7.31	2.54

The *PA*-list was composed of 14 pairs of adjectives of low similarity taken from Melton's lists.[2] For every *S*, half the pairs served as experimental items and were taken from the 14-item *Sr*-list previously learned. The arrangement of the items in the experimental pairs was identical to the ordering of the items in the prior *Sr*-list. The other half of the pairs served as control pairs. To control for differential difficulty due to intrinsic differences in ease of learning, one half the *S*s in each group learned one *Sr*-list while the other half learned the second *Sr*-list. The effect of this procedure is to treat each half of the pairs in the *PA*-list as experimental pairs for half the *S*s and as control pairs for the other half of the *S*s. The design of the experiment may then be described as a 2 × 3 × 2 mixed factorial design with 2 *Sr*-lists, 3 rates of *PA*-presentation, and 2 sublists of *PA*-pairs being employed.

The lists were presented by a film-strip projector with approximately 5 min. between the learning of the *Sr*- and the *PA*-lists to allow time to change film-strips and to read instructions. The criterion of learning for all lists was one perfect recitation. Three orders of presentation were used for the *PA*-list to reduce the possibility of *Sr*-learning.

Results. Initially it should be noted that *PA*-learning was a function of rate of *PA*-presentation. The means of trials to learn the *PA*-list for each of the three groups are presented in Table I. These data were found to be heterogeneous and a logarithmic transformation was employed. The analysis of the transformed data indicates that speed of learning increased as rate of presentation decreased ($F = 12.58$, $df = 2/42$, $P < 0.01$).

---

[2] E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951, 548-552.

The means of trials to learn the sublists in the *PA*-list were analyzed to test the hypothesis proposed by the *S*s. In this analysis rate of learning increased as rate of presentation decreased ($F = 9.45$, $df = 1/42$, $P < 0.01$) and the two seven-item sublists of the *PA*-list were not of equivalent difficulty ($F = 12.20$, $df = 1/42$, $P < 0.01$). No differences were found between the learning of the experimental and control sublists ($F = 2.62$) or in the interaction between the learning of these conditions and rate of presentation of the *PA*-list ($F < 1.00$).

The means of correct responses per trial during *PA*-learning were analyzed. The means and variances of these data were correlated and the square root transformation was employed. In this third analysis rate of presentation was also found to be a significant variable ($F = 11.77$, $df = 2/42$, $P < 0.01$) reflecting the results of the analyses of trials to learn. In addition the two sublists of the *PA*-list differed in difficulty ($F = 13.80$, $df = 1/42$, $P < 0.01$).

The comparison of the experimental and control sublists in the *PA*-list provided no evidence supporting the introspective hypothesis of the *S*s. The means of correct responses per trial tended to be greater for the control pairs than for the experimental pairs, but this apparent negative transfer ($F = 3.10$, $df = 1/42$, $P < 0.10$) failed to reach the 0.05 level of significance. In addition, the interaction between transfer (*i.e.* experimental vs. control) and rate of *PA*-presentation was not significant ($F < 1.00$), also failing to support the hypothesis that positive transfer should increase as rate of presentation decreased.

Although no evidence supporting the hypothesis of the *S*s has been found in the analysis of the learning of the *PA*-list, an analysis of the learning of individual items within the *PA*-list may provide evidence favoring such an interpretation and give additional information about the apparent negative transfer obtained. As expressed by the *S*s, the *Sr*-list is 'gone through' during *PA*-learning and it would appear that the *S* would start at the beginning of the *Sr*-list and proceed toward the end each time a *PA*-stimulus was presented. If such were the case, then the pairs constructed from items at the beginning of the *Sr*-list would have been correct more often relative to their control than would have been the pairs constructed from items taken from the end of the *Sr*-list.

Every *S* learned one of two *Sr*-lists and the appropriate control for a specific pair is that pair which would have occupied the identical serial position had the other *Sr*-list been learned. Selection of appropriate control items is discussed to a greater extent elsewhere.[3] The means of correct responses per trial during *PA*-learning were computed for each serial position and the means of the respective control pairs were subtracted from the means of the experimental pairs. These data are presented in Fig. 1 with the ordering of the data along the baseline identical to the ordering of the items in the *Sr*-list. As can be seen from the figure, the last items were the only items to be learned faster than the control and these items are those least expected by the hypothesis of the *S*s to be learned faster. The learning of each pair was compared to its control pair and three experimental pairs were found to differ from their controls. The pairs in positions 4 ($F = 8.64$, $P < 0.01$) and 5 ($F = 6.00$, $P < 0.05$) were learned slower than their controls while the pair in position 7($F = 24.30$, $P < 0.001$) was learned faster.

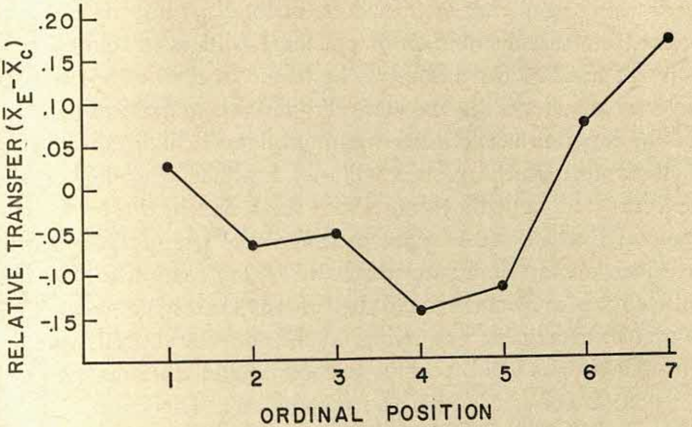


FIG. 1. RELATIVE TRANSFER IN THE LEARNING OF THE *PA*-LIST WITH THE ORDER OF THE PAIRS THE SAME AS IN THE *Sr*-LIST

The shape of the curve presented in Fig. 1 was consistent over the three conditions of rate of *PA*-presentation and the curve, when analyzed for trend, was found to have a significant quadratic component ($F = 26.33$, $df = 1/42$, $P < 0.001$).

Discussion. As rate of presentation decreases, rate of *PA*-learning increases. This finding is consistent with other research in which rate of learning was found to be inversely related to rate of presentation.[4] It should be noted that intertrial interval increased as rate of presentation decreased and the increase in rate of learning might be interpreted in terms of differences due to distribution of practice. Aside from the fact that a change in the intertrial interval from 3 sec. to 8 sec. does not seem to be of a magnitude which would influence rate of learning, there is some evidence that rate of learning would *not* be a function of intertrial interval. Underwood has pointed out that distribution of practice facilitates learning only when the items of a task must themselves be learned (as in experiments using nonsense syllables).[5] When the items of a task are available to the *S* at the start of the experiment (i.e. when meaningfulness is high as in the present experiment) distributed practice will not facilitate learning. Hence the differences in rate of *PA*-learning which are found in the present experiment appear interpretable in terms of differential rate of *PA*-presentation.

---

[3] Young, *op. cit., J. exp. Psychol.*, 63, 1962, 307-313.

[4] C. I. Hovland, Experimental studies in rote-learning theory: VIII. Distributed practice of paired associates with varying rates of presentation, *J. exp. Psychol.*, 39, 1949, 714-718.

No evidence was found supporting the *S*s' explanation for the absence of positive transfer from the *Sr*- to the *PA*-list. Their hypothesis was tested in the present experiment by varying rate of presentation of the *PA*-list. If their hypothesis was valid, positive transfer should have increased as rate of presentation decreased. Such was not the case. There was also no evidence for positive transfer as would be predicted by the specificity hypothesis.[6] This hypothesis assumes that the nominal stimulus in *Sr*-learning is the functional stimulus and as such positive transfer should obtain from *Sr*- to *PA*-learning.

Rather than positive transfer obtaining from the *Sr*- to the *PA*-list, some evidence was found indicating negative transfer had occurred. Because of this, the transfer of specific items within the list was investigated, and the pair constructed from the last two items of the *Sr*-list was learned faster than its control while two pairs taken from the middle of the *Sr*-list were learned slower than their controls. Additional analysis indicated that the quadratic component of the data presented in Fig. 1 was significant. This quadratic component might be interpreted in accordance with predictions from the specificity-hypothesis as transfer of the serial-position curve from the *Sr*- list to the *PA*-list. In contrast to predictions from the specificity-hypothesis, however, little positive transfer was observed and the pair which should have shown maximal positive transfer (the pair taken from the initial *Sr*-positions) shows no transfer at all.

While the predictions derived from the hypothesis of the *S*s have not been supported, the *S*s still may have attempted to learn the *PA*-list in the manner they described. If they did attempt to learn the pairs in the experimental sublist by recalling the previously learned *Sr*-list, it would then be expected (because of generalization of serial position) that the recall of the items from the middle of the *Sr*-list would be poorer than recall of the items at the ends.[7] Furthermore, if the *S*s' method did *not* facilitate the learning of the experimental pairs but was continued for several trials, then the middle items might show negative transfer when compared to their controls. The difficulty with this explanation is that the positive transfer shown by the last pair can not be accounted for.

---

[5] B. J. Underwood, Ten years of massed practice on distributed practice, *Psychol. Rev.*, 68, 1961, 229-247.

[6] Young, *op. cit.*, this JOURNAL, 74, 1961, 517-528.

Another explanation of the results of the present experiment may be based on research concerned with the effective stimulus in *Sr*-learning. Previous research has given support to the idea that associations are formed between items at the ends of the *Sr*-list (the specificity-hypothesis) while in the middle of the *Sr*-list associations are formed between the position of the item and the item itself (the serial-position-hypothesis).[8] Position does not appear to be a stimulus in the *PA*-learning because of the varied order of presentation and thus the results predicted for the present experiment would be that the ends of the *PA*-list would be learned faster than the middle. Although the first item, in contrast to predictions, was not learned faster than its control, it was learned faster than the items in the middle of the list and the quadratic component of the transfer curve could be accounted for in this manner.

Summary. The present experiment tested the hypothesis that absence of transfer from a serial (*Sr*) list to a paired-associate (*PA*) list was due to the rapid rate of *PA*-presentation. The hypothesis was tested in a 'mixed-list' experimental design in which each *S* served as his own control. Each *S* first learned an *Sr*-list with a 1.5-sec. rate of presentation and then learned a *PA*-list in which half the pairs were taken from the previously learned *Sr*-list. The independent variable was rate of *PA*-presentation which varied from a 1.5-sec. rate of presentation (1.5 sec. for the stimulus alone and then 1.5 sec. for the stimulus and response together) to a 4-sec. rate of presentation. Contrary to predictions from the hypothesis, positive transfer was not observed and did not increase as rate of *PA*-presentation decreased. Analysis of individual items within the *PA*-list indicated that pairs taken from items in the middle of the *Sr*-list were learned slower than the control, while pairs taken from items at the end of the *Sr*-list were learned faster than the control.

---

[7] R. W. Schulz, Generalization of serial position in rote serial learning, *J. exp. Psychol.*, 49, 1955, 267-272.

[8] Young, *op. cit.*, *J. exp. Psychol.*, 63, 1962, 307-313.

THE INFLUENCE OF CONTEXT UPON THE ESTIMATION OF NUMBER

By WILLIAM BEVAN, RICHARD A. MAIER, and HARRY HELSON,
Kansas State University

The principle of adaptation-level states that judgments of magnitude are the product of both focal and contextual stimuli. This has been demonstrated for a variety of sensory dimensions and other dimensions, as well.[1] The present experiment deals with a judgmental task of the more cognitive sort—the estimation of number—under conditions in which the context of the stimuli was varied. Actually, two properties of the context—size and figure-ground relation—were selected for experimental manipulation. The former is identified with a physical property of the stimulus-field. It is analogous to the weight of anchors outside the series in the weight-lifting experiment and the lightness of the background in the color-conversion experiment.[2] The latter refers to changes in the organization of the perceptual field with the stimulus held constant. It is concerned with the influence of contextual stimuli as a function of their incorporation into the figure or the ground. The phenomenon of appurtenance, as well as Gelb and Granit's demonstration of a difference in visual threshold when the stimulus-patch fell in a figure-area, as opposed to a ground-area, suggests the importance of this second variable.[3]

The experimental task consisted of making direct estimates of the number of beans in each of several glass jars by four groups of *S*s. For two groups, the jars were large and were identified in one case as figure, in the other as ground. For the two remaining groups, these conditions were repeated with small jars. It was anticipated that jar-size would exert a significant influence upon average estimates of number, and that this effect would be most pronounced when the jar, along with the beans, was identfied as *figure.*

---

* Received for publication February 10, 1962. This work was supported in part by the Office of Naval Research under terms of Contract Nonr-3634(01).

[1] Harry Helson, Adaptation-level theory, in Sigmund Koch (ed.), *Psychology: A Study of a Science,* 1, 1959, 565-621.

[2] Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29; Fundamental problems in color vision: I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination, *J. exp. Psychol.,* 23, 1938, 439-476.

[3] Kurt Koffka, *Principles of Gestalt Psychology,* 1935, 187-188, 246.

Method: (1) Subjects. In all, 300 men, 75 per group, made the estimates. All were students in introductory psychology and were told nothing about the purpose of the experiment except that it was concerned with their ability to make estimates in terms of numbers.

Materials. The beans were common red dried beans, averaging about one centimeter in length. Each group of *S*s gave estimates of three samples, 60, 230, and 410 beans, all in jars of one size. The jars were cylindrical in shape, the small ones being approximately 2 in. in diameter and 4.5 in. high; the large jars were 3.5 in. in diameter and 7.25 in. high. In the small jars, the beans came to $\frac{1}{9}$, $\frac{1}{3}$, and $\frac{2}{3}$ of the total height of the bottle; in the large jars, they came to $\frac{1}{28}$, $\frac{3}{28}$, and $\frac{1}{7}$ the total height.

(2) Procedure. The *S*s were tested in small groups. They were asked to gather about a table where the jars were to be displayed, handed small sheets of paper on which to write their estimates, and then given the following instructions:

> *Instructions.* This is a little study of number-estimation. I am about to show you three jars, each containing a different number of beans. Number 1 contains the smallest number of beans; Number 2, a medium number; and Number 3 the largest number. When I place the jars before you, please write down, *without counting, your first impression* of the number of beans in each.

Manipulation of the variable of field-articulation, it is assumed, was successfully achieved through differential instructions. For two groups, the integral relationship between beans and jars as figure was stressed, while for the other two, the beans as figure were contrasted to the jars as ground. Thus, the 'figure' and 'ground' instructions were respectively: "One thing to keep in mind is that beans and jar go together, constituting, as it were, a single organic unit set off from their surroundings," and, "While it may be natural for you to think of the beans and jars as one, remember that they are separate. The jar is a container or receptacle and can be thought of as part of the context or background, just as in this picture of the goblet the black areas are the background for the white vase." At this point *E* displayed the Rubin goblet-figure, and then took the three jars to be judged from a paper sack and placed them before the *S*s.

Results: (1) Direct estimates of number. Examination of Fig. 1 indicates immediately that all samples of beans were grossly underestimated. The 60-bean sample was, on the average, underestimated by 37%, the 240-bean sample by 45%, and the 410-bean sample by 42%.

This finding agrees with earlier findings by Jenness, who had *S*s make direct estimates of a single sample of 811 beans under a variety of conditions.[4] There was a slight over-all tendency toward underestimation on the part of the men studied by Jenness, and a more marked tendency toward overestimation on the part of his women. After group-discussion, both sexes displayed a marked tendency toward underestimation. Taves has reported that when the dimension of numerousness is scaled by the method of fractionation, using samples of from 2 to 180 light dots on a dark background, the half-judgments are objectively correct up to 8 or 10 dots.[5] Beyond this number, there is a tendency toward overestimation. When Taves had *Ss* make direct estimates, however, he found the tendency to overestimate with sample sizes up to 60, with a tendency toward underestimation of larger samples.

---

[4] Arthur Jenness, The role of discussion in changing opinion regarding a matter of fact, *J. abnorm. soc. Psychol.*, 27, 1932, 279-296.

Downward extrapolation in the case of the present data suggests that the curves converge toward correct estimates (dotted line of Fig. 1) at about 20 beans. Taken altogether, the results in the literature suggest the direction of the constant error in estimations of numerousness to depend to significant degree both upon sex and upon the psychophysical method employed.

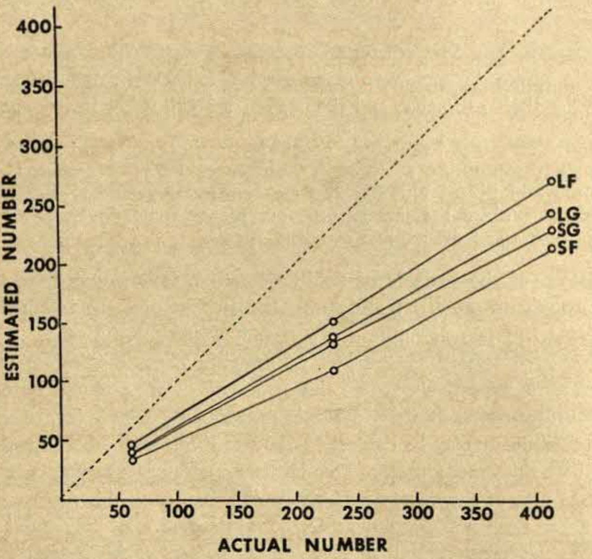


FIG. 1. DIRECT ESTIMATES OF NUMBER AS A FUNCTION OF JAR-SIZE AND FIELD-ARTICULATION

L and *S* indicate large and small jars, and *F* and *G* the identification of the jar as figure and ground, respectively. The dotted line indicates where the judgments would fall if the *Ss* were perfectly accurate.

(2) Influence of jar-size upon estimates. The data of Fig. 1 show that while the number of beans was underestimated in all four conditions, the estimates were consistently larger with the large containers than with the small ones ($F = 6.37$, $df = 1$ and 296, $p < 0.02$). Furthermore, the differences in judgment associated with jar-size increased in the size of

---

[5] E. H. Taves, Two mechanisms for the perception of visual numerousness, *Arch. Psychol.,* 37, 1941, (No. 265), 1-47.

the sample estimated. It would thus appear that numerical estimates, like sensory magnitudes, reflect the influence of context.

There are a variety of data which indicate that stimulus-properties on one dimension influence apparent values on others. There is, for example, the commonly recognized fact that loudness varies with frequency as well as amplitude. Similarly, Burnham has shown that color varies with patch-size, while Bevan and Dukes have reported the converse relationship.[6] This sort of dimensional interaction has been reported for numerousness. Ansbacher discussed data of Kasting which indicated that estimations of number vary with stimulus-size, and Ansbacher's own data indicated the same relationship for apparent number and monetary value.[7]

Just as judgment reflects significant relationships among stimulus-dimensions, it also reflects stimulus-background interactions. One need cite only Helson's experiments on color-conversion or the Gestalt studies on orientation and apparent shape to provide ample evidence.[8] The dimension of numerousness also demonstrates the dependence of figural-property upon context. Mokre has reported an inverse relationship between numerousness and the spatial dispersion of dots.[9]

The present data would appear to be inconsistent with this finding, for, although they are marked by a definite trend toward underestimation, increasing the volume of the contextual stimulus results in a reduction in this trend. Meanwhile, it must be noted that the two experiments differ not only in the materials but also in the psychophysical method used. Furthermore, that increasing the spatial dispersion of the focal stimulus, on the one hand, and the size of the contextual stimulus, on the other, are equivalent operations may not be assumed.

(3) Influence of field-articulation upon estimates. The data of Fig. 1 provide support for the prediction that jar-size exerts a greater influence upon judgment under the 'figure' attitude than under the 'ground' attitude. As seen from Fig. 1, size of jar as figure exercises a definite effect on the estimates in pushing the estimates up when the large one is figure and in pushing them down when the small one is figure. As ground, size of the containers has little effect. Estimates made under the 'ground' instructions differ hardly at all from the over-all average for all conditions. In contrast, the two curves associated with the 'figure' instructions represent definite and complementary departures from the over-all average. This

---

[6] R. W. Burnham, The dependence of color upon area, this JOURNAL, 64, 1951, 521-533; William Bevan and W. F. Dukes, Color as a variable in the judgment of size, this JOURNAL, 66, 1953, 283-288.

[7] Heinz Ansbacher, Perception of number as effected by the monetary value of objects, *Arch. Psychol.*, 30, 1937 (No. 215), 1-88.

[8] Helson, *op. cit.*, *J. exp. Psychol.* 23, 1938, 439-476; Harry Helson and V. B. Jeffers Fundamental problems in color vision: II. Hue, lightness, and saturation of selected samples in chromatic illumination, *ibid.*, 26, 1940, 1-27.

[9] H. Mokre, Über den Einfluss von Grösse und Abstand der Elemente auf die Mengenaffassung, *Z. Psychol.*, 105, 1928, 195-225.

is supported by a significant size × instructions interaction when the one-tailed criterion is applied ($F = 2.82$, $df = 1$ and 296, $p < 0.05$).

The functional difference between figure and ground is also demonstrated by the results of Gelb and Granit.[10] These investigators measured the visual threshold for a colored dot of light when it fell on figure-areas as opposed to grounds, with brightness-differences counterbalanced. The threshold was consistently lower when on ground than when on figure.

Wide individual differences in the impression of numerousness are reflected in the fact that individual estimates for particular conditions differed by orders of magnitude of as much as 1-10. Taves reported the same ratio of differences for his direct estimate.[11]

(4) Context-effects and semantic relativity. Shifts in judgmental value with changes in context are explained by the principle of adaptation-level in terms of labile internal norms that are the product of a combination of inputs from both the present judgmental situation and the judge himself. The apparent magnitude of the focal stimulus, *i.e.* the stimulus to be judged, is conceived to be relative to the norm and thus subject to change with a change in the norm. Opposed to this view is the Fechnerian one, which defines the psychophysical relationship in terms of the intensitive properties of the focal stimulus alone. Reflecting this latter tradition, Stevens has had to maintain that, outside the domain of sensory adaptation, shifts in scale-value can reflect only adjustments of the verbal labels used to characterize the focal stimulus.[12] Thus, for example, when an anchor is added above a series of lifted-weights, the values assigned to the series-stimuli must be adjusted downward in order to reflect the increased range of inputs. Meanwhile, the apparent magnitude of the stimulus is assumed not to shift, only the language changes. In instances involving purely verbal materials (*e.g.* the study by Helson, Dworkin, and Michels of the quantitative denotation of common terms) the notion of semantic relativity appears plausible.[13] Where, however, judgment concerns physical magnitudes and intensities, and certain qualities like hardness and redness, direct experience indicates that the net effect of stimulation depends upon contextual stimuli.

A preponderant number of experiments on context-effects have involved the use of rating scales and therefore are potentiaally vulnerable to Stevens' argument; but the subjects of the present experiment were required to render their judgments of number in the system of natural integers. Such a judgmental language is absolute and maximally stimulus-bound. When he estimates that there are 150 beans in Jar 2, *S* does not mean covertly or overtly that 150 is small relative to the number of beans in Jar 3 or large relative to the number in Jar 1, or that his estimate has any other meaning except that associated with the system of numbers. Along with the great American mathematician, George Birkhoff, who, holding up five fingers, once remarked, "Whenever I find a collection of objects that exactly matches the fingers of my hand, I will call that collection five," he is saying, "This is what I mean by 150." Indeed, he is performing the operation of ostensive definition which Stevens, himself, has stressed as essential to the scientific study of behavior.[14] The argument of semantic relativity can thus have no validity for the contextual effects found in this study.

---

[10] Koffka, *op. cit.,* 187-188.

[11] Taves, *op. cit.,* 35.

[12] S. S. Stevens, Adaptation-level vs. the relativity of judgment, this JOURNAL, 71, 1958, 633-646.

[13] Harry Helson, R. S. Dworkin, and W. C. Michels, Quantitative denotations of common terms as a function of background, this JOURNAL, 69, 1956, 194-208.

SUMMARY

This experiment investigated the role of context in the estimation of number. Two variables, size and figure-ground relationship, were manipulated. Four groups of men, 75 per group, gave direct estimates of three samples of beans when the beans were presented either in large or in small jars and when the jars, through instructions, either were identified along with the beans as figure or were differentiated from the beans as ground.

There was a consistent and marked tendency to underestimate the numbers in all three samples. Size of jar as figure exercised a definite effect on the estimates in pushing the estimates up when the large one was figure and in pushing them down when the small one was figure. As ground, size of the container had little effect. Finally, contrary to the point of view of semantic relativism, these context-effects were demonstrated with an absolute judgmental language.

---

[14] Stevens, The operational definition of psychological concepts, *Psychol. Rev.*, 42, 1935, 517-527.

BIDIRECTIONAL PAIRED-ASSOCIATE LEARNING

By BENTON J. UNDERWOOD and GEOFFREY KEPPEL,
Northwestern University

In the usual verbal paired-associate task one set of items serves as stimulus-terms and another set as response-terms; these functions are constant throughout learning. The measures of performance, therefore, commonly include only unidirectional associations, *i.e.* the frequency with which the stimulus-terms elicit the paired response-terms. As a number of studies have shown, however (*e.g.* Richardson[1]), in the process of learning the forward, or S–R, associations, backward, or R–S, associations also develop. While the strength or frequency of these backward associations is seldom as great as the forward associations, it is clear that bidirectional associations do develop during the usual paired-associate procedure.

The present study examines paired-associate learning when bidirectional learning is *required* of the *S*. Thus, instead of always presenting a pair of items in a constant order, A-B, as is done typically, the order is sometimes A-B and sometimes B-A. *S* is simply instructed to give the other member of the pair, regardless of which member is presented as the stimulus-term. To assess the effect of requiring bidirectional learning, Control *S*s must be given the typical or unidirectional paired-associate task.

There are at least three factors whose influence would predict that bidirectional learning would require more time than unidirectional learning. First, in bidirectional learning, *S* would have to learn twice as many response-terms as would an *S* learning unidirectionally. In the present experiment, the differential influence of this factor for the two learning procedures has been minimized by the use of high-frequency words. Secondly, in bidirectional learning *S* cannot classify one set of items as stimulus-terms and another set as response-terms. While such a classification may occur in unindirectional learning, it is not known to what extent it occurs, and, if it does occur, whether or not it influences rate of learning. Such classification might be expected, however, to facilitate learning on the grounds that the number of potential interfering tendencies is reduced thereby. That is, if one set

---

* Received for publication September 18, 1961. This work was done under Project 154-057, Contract Nonr-1228 (15) between the Office of Naval Research and Northwestern University.

[1] Jack Richardson, Comparison of S-R and R-S learning of paired-associates, *Psychol. Rep.*, 7, 1960, 225-228.

of items can be classed as stimulus-items and another set as response-items, the interference between classes may be less than the interference which would prevail if classification did not occur. In the present study, an attempt has been made to minimize differential effects of interference for bidirectional and unidirectonal learning by using words with low meaningful and formal similarity.

The third reason why bidirectional learning may be expected to require more time than unidirectional learning relates directly to backward learning. It has been noted that backward learning is usually not as complete as forward learning following unidirectional learning procedures. (Why *any* backward learning occurs is still not known, and why it is less than forward learning is also not known, but neither issue is critical for the present study.) It would seem, therefore, that if *S* is required to learn both the forward and backward associations, and to learn them both to the same level as forward associations under unidirectional procedures, bidirectional learning should take more time than unidirectional learning. Indeed, if our analysis of the situation as given above is complete, the time required to complete bidirectional learning would be predicted to consist of the time required for unidirectional learning plus the time necessary to bring the strength of backward associations in unidirectional learning to the level of strength of forward associations. This simple summation assumes that differences in response-learning and differences in the effects of interference are minimal for bidirectional and unidirectional procedures.

METHOD

Lists. The materials consisted of 10 pairs of two-syllable adjectives, as follows: *clever-sacred, active-distant, open-weary, forward-royal, entire-faithful, supreme-liquid, idle-angry, dirty-central, healthy-evil, merry-little.* Nine of the words have a Thorndike-Lorge frequency-count of 27 to 50 per million, eight have *A* frequency-ratings, and three have *AA* ratings.[2] Three different lists were made from the 10 pairs. One list was the bidirectional list learned by the Experimental *S*s. A second list was a unidirectional list which will be called A-B, this list being learned by half the Control *S*s. The third list was another unidirectional list which consisted of a 'turned over' version of A-B, to be called B-A; this list was learned by the other half of the Control *S*s. The list as given above constitutes the A-B list; when *sacred, distant, weary,* etc., were stimulus-terms, the result was the B-A list.

The lists were presented on a memory drum at a 2:2-sec. rate. Four different orders of the pairs were used to prevent serial learning, these orders being the same for all *S*s. For the bidirectional list, a given word of a pair served as a stimulus-term in two of the four orders and as a response-term in the other two of the four orders. This change in position, however, was nonsystematic from order to order when all pairs were considered.

Procedure and subjects. The Experimental Group consisted of 30 *S*s, and each of the Control Groups had 15 *S*s. All *S*s were college students, most of whom were practiced in unidirectional paired-associate learning. All three groups learned to a criterion of one perfect trial by the anticipation-method. The Experimental *S*s were fully informed of the nature of the bidirectional list before it was presented. Immediately after learning, both Control Groups were given a backward-learning test. This test consisted of presenting the B-A list to the Control *S*s who had just learned the A-B list, and presenting the A-B list to the Control *S*s who had just learned the B-A list. The *S*s were fully informed concerning the nature of the 'new' list and were asked to give as many correct responses as possible on the first trial. Learning was then carried to one perfect trial, but never less than five trials.

---

[2] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words,* 1944.

After the 30 *S*s in the Experimental Group had learned the bidirectional list, they were told that additional trials would be given but that the position of the items would be fixed, *i.e.* a given member of a pair would always occur as the stimulus-term. For half the *S*s the fixed list was the A-B list, and for the other half, the fixed list was the B-A list. The learning of this fixed list was carried to one perfect trial, but never less than five trials.

The 60 *S*s were assigned to the conditions on a random basis. Whenever an *S* was lost for any reason, he was replaced in the schedule of conditions by the next *S* appearing at the laboratory. Two *S*s were lost for failure to learn within the 50-min. period, one for failure to learn the bidirectional list, and one for failure to learn a unidirectional list.

RESULTS AND DISCUSSION

The 30 *S*s in the Control Group required a mean of 17.10 trials to learn to one perfect trial. The 30 *S*s in the Experimental Group required a mean of 22.07 trials. The difference (4.97 ± 2.88) produces a *t* of 1.73, a value which is short of the 5% level of significance. It would seem, therefore, that on strict statistical grounds, bidirectional learning does not require more time than unidirectional learning for the present materials. Considerable other evidence suggests, however, that there is a reliable difference between the two forms of learning in spite of the fact that the trials-to-criterion measure does not reflect this strongly.

If bidirectional associations are learned as quickly as unidirectional associations, it would be expected that following the first correct anticipation for a given pair, the probability of a correct anticipation on the immediately following trial for this pair would be equivalent for Experimental and Control *S*s. These probabilities are, in fact, not equivalent. For each *S* we examined the records for each pair to determine whether or not it was given correctly on the trial immediately following the trial on which it was first correctly anticipated. The positive instances were summed for each *S* and the mean for all *S*s in each group determined. For the 30 Control *S*s, this mean was 6.03, but for the Experimental *S*s, the value was only 4.53. The difference (1.50 ± 0.48) gives a *t* of 3.13. Clearly, a correct anticipation of a response under unidirectional learning is associated with a greater probability that that item will be given correctly on the immediately following trial than is true under bidirectional learning. Why should this difference occur? Our interpretation is that when a given item is correctly anticipated the first time, the forward association is stronger than the backward association. Therefore, since the position of the two words of some of the pairs is changed on the immediately following trial under bidirectional learning, the correct anticipation depends in part, at least, upon the strength of the backward association developed previously. This conclusion need not, however, necessarily follow. An *S* in the Experimental Group *could* learn all or most items unidirectionally and still attain a perfect recitation. This could be accomplished by learning the items unidirectionally as determined by one order of presentation and ignoring the items when the order of the members of a pair changed. If this did happen, it would be apparent in cyclical performances on the data-sheets; however, evidence for such an effect is lacking in the data-sheets. Furthermore, the data on transfer, to which we turn now, make it clear that most Experimental *S*s learned bidirectionally.

On the first transfer-trial, the Experimental Group gave 8.53 correct responses, the Control Group gave 6.40. The difference (2.13 ± 0.47) gives a *t* of 4.53. The fact that the Control Group recalled an average of 6.40 items indicates that backward learning occurred during original unidirectional learning. But the fact that the Experimental *S*s recalled more than the Control *S*s indicates that the associations in *either* direction for the Experimental *S*s were stronger than the backward associations of the Control *S*s. In effect, then, the Experimental *S*s learned more than the Control *S*s. This in turn, may indicate that the greater number of trials required by the Experimental *S*s to learn originally is not a chance phenomenon, even though the difference in trials to learn did not reach the 5% level of significance. The superiority of the Experimental over the Control Groups on transfer is further shown in the 'relearning' scores. The Experimental *S*s required a mean of 3.07 trials, the Control *S*s, 8.40 trials. The difference (5.33 ± 1.29) gives a *t* of 4.13.

In the introductory section, it was argued that for the present materials the only difference between bidirectional and unidirectional learning would be that in unidirectional learning *S* is not *required* to learn what for him would be backward associations. The data may now be viewed with this notion in mind. The Experimental *S*s required 22.07 trials to learn, the Control *S*s required 17.10 trials. Thus, the Control *S*s required 23% fewer trials to learn than did the Experimental. On the first transfer-trial the Experimental *S*s, having learned bidirectionally, gave 8.53 items correctly, while the Control *S*s gave 6.40 correct responses. This latter value represents the number of backward associations acquired during unidirectional learning. Thus, the Control *S*s recalled 25% fewer items than did the Experimental. The difference in time to learn (23%) is almost perfectly matched by the difference in degree of learning (or number) of backward associations (25%). The data may be viewed in still another way. At the end of original learning the Experimental *S*s had presumably acquired bidirectional associations. The equivalent amount of learning was attained by the Control *S*s only after having learned the 'turned-over list' on the test of transfer. If the criterial trials are not counted, it can be seen that the Control *S*s learned bidirectional associations after 16.10 plus 7.40 trials, or 23.50 trials. The Experimental *S*s reached the same level of learning in 21.07 trials.

These data suggest that in bidirectional learning with the present materials no unique phenomena are introduced; we are simply measuring the time required to acquire forward and backward associations simultaneously and this acquisition time is approximately equal to the time required to learn forward associations followed by the learning or strengthening of the backward associations for the same items.

SUMMARY

This experiment compared bidirectional paired-associate learning, wherein either member of a pair of words may occur as the stimulus-term or as the response-term, with typical unidirectional learning, where a word is always a stimulus-term or always a response-term but never both. A total of 30 *S*s learned the bidirectional list and 30 learned a unidirectional list as a control. Tests for backward associations developed during unidirectional learning were made for the latter control *S*s.

Bidirectional learning required more time than did unidirectional learning. Bidirectional learning may be interpreted as representing a simple summation of time to learn forward and backward associations to the same level of strength.

THE RELATIONSHIP OF LATERALITY AND CUTANEOUS AREA TO BREAST-SENSITIVITY IN SINISTRALS AND DEXTRALS

By SIDNEY WEINSTEIN, Albert Einstein College of Medicine

In a recent study, superiority of the left hand over the right for punctate pressure-sensitivity was found in a group of normal dextrals.[1] Subsequent work has confirmed the relative superiority of the nondominant upper extremity in a predominantly dextral group of children,[2] and in a group of dextral adults.[3] Both in the children and in the adults, the arms tended to show less difference than the palms, which suggested that motility of bodily parts may predispose to the development of differential sensitivity. The opportunity to test this hypothesis arose in the course of a larger investigation of the somatosensory after-effects of mastectomy. Punctate pressure-sensitivity and two-point discrimination were examined as a function of laterality, handedness, and breast-size.

In all previous studies of breast-sensitivity, factors of laterality and handedness have been ignored, and attention has been concentrated almost exclusively upon the nipples and the areola.[4] Belonoschkin, who tested the nipples of women prior to mastectomy for cancer, observed that sensitivity to warmth, pressure, and pain was quite poor.[5] Goldscheider likewise found the areola and nipple to rank extremely low in relation to other bodily parts for sensitivity to warmth and cold.[6] Wood-Jones and Turner described the areola and nipple as "typically protopathic" tissues and quite insensitive to light sensory stimuli produced by hairs or

---

* Received for publication December 3, 1961. This work was supported by Grants H-3838, M-2573, A-2965, and NIHB-3356 from the National Institutes of Health. The author is indebted to Robert Wenk, Robert Vetter, and E. A. Sersen for aiding in the statistical analyses.

[1] Josephine Semmes, Sidney Weinstein, Lila Ghent, and H.-L. Teuber, *Somatosensory Changes After Penetrating Brain Wounds in Man*, 1960, 6.

[2] Lila Ghent, Developmental changes in tactual thresholds on dominant and nondominant sides, *J. comp. physiol. Psychol.*, 54, 1961, 670-673.

[3] Sidney Weinstein and E. A. Sersen, Tactual sensitivity as a function of handedness and laterality, *J. comp. physiol. Psychol.*, 54, 1961, 665-669.

[4] F. Wood-Jones and J. B. Turner, A note on the sensory characters of the nipple and areola, *Med. J. Australia*, 1, 1931, 778-779; C. S. Sherrington, Cutaneous sensation, in E. A. Schaefer (ed.), *Textbook of Physiology*, 2, 1900, 949, 961; Boris Belonoschkin, Physiologisch-Anatomische Untersuchungen über die Empfanger der Kaltemfindung, *Z. Zellforsch, u. mikr. Anat.*, 18, 1933, 555-572.

[5] Belonoschkin, *op. cit.*, 555-572.

[6] Sherrington, *op. cit.*, 949, 961.

sandpaper.[7] Accordingly, it was decided to avoid the areola in the present study and to determine instead the sensitivity of the surrounding breast.

Method: (1) Subjects. The *S*s were 16 dextral and 16 sinistral women. The median age (and range) of the dextrals was 19.5 (17-35) yr., and that of the sinistrals was 19.0 (17-25) yr. All *S*s were normal volunteers from local colleges. Since previous work has shown sensitivity to vary with sex-hormone concentration,[8] the *S*s were questioned to determine whether they were being tested during the premenstrual, menstrual, postmenstrual, or intermenstrual parts of their cycles. The numbers of dextral *S*s within the different parts of the cycle were 3, 3, 3, and 7 respectively; the numbers of sinistral *S*s within the same parts of the cycle were 2, 2, 4, and 8.

(2) Determination of handedness. All *S*s completed the questionary currently being used in this laboratory dealing with hand-preference for a number of items. All the sinistrals considered themselves to be sinistral, and wrote with their left hands. The converse was true of all dextrals.

(3) Design. For half of each group, the right breast was tested first and, for the other half, the left was tested first. In both groups, and for each order of testing with regard to laterality, equal numbers of *S*s were tested first for pressure-sensitivity and first for two-point discrimination.

(4) Apparatus and testing. Punctate pressure-sensitivity was measured by means of a pressure-esthesiometer.[9] Two-point discrimination was measured by means of a two-point esthesiometer.[10] Both thresholds were determined by means of three alternating ascending and descending determinations in a modified method of limits.[11] *S* sat erect in a chair during measurement and testing. Her eyes were closed during testing.

(a) Locus of testing. Punctate pressure-sensitivity was tested at a point 6 mm. above the rostral border of the areola. Two-point discrimination was tested transversally, with the two points straddling the region used for the test of pressure-sensitivity.

(5) Measurement of area. Four measurements were made on each breast. One was the circumference of the base of the breast; the second was the circumference of the breast measured 6 mm. distal from the areolar border at all points; the third was the distance from the center of the nipple to the base of the breast on the ventral surface. The product of this altitude and the mean of the two circumferences yielded an approximate measurement of the total cutaneous area of each

---

[7] Wood-Jones and Turner, *op. cit.,* 778-779.

[8] R. Y. Herren, The effect of high and low female sex hormone concentration on the two-point threshold of pain and touch and upon tactile sensitivity, *J. exp. Psychol.,* 16, 1933, 324-327.

[9] The instrument and method of testing are described in Semmes, Weinstein, Ghent, and Teuber, *op. cit.,* 5, 59-62. It is manufactured by Shaw Laboratories, 21 Harriet Drive, Syosset, L.I., N.Y.

[10] The two-point esthesiometer is adapted from machinist's calipers and calibrated in mm. The hemispherical area of the point of contact of the single brass point is equal to the total area of the contact-surfaces of the two brass points. The esthesiometer is manufactured by Shaw Laboratories, 21 Harriet Drive, Syosset, L.I., N.Y.

[11] The method of testing is described in Semmes, Weinstein, Ghent, and Teuber, *op. cit.,* 6-7.

breast. The fourth measure was the distance between the center of the nipple and the point tested, *i.e.* areolar radius plus 6 mm.

Results. The groups first were compared in terms of breast-area and areolar radius. A two-factor analysis of variance evaluated the effects of laterality and handedness for each measure separately. There were no differences between the breasts for total area or for areolar radius, nor were there differences between the sinistral and dextral groups for either measure on either breast. Pearson correlations between the breasts for total area yielded *r*s of 0.93 for the dextrals and 0.94 for the sinistrals.

An analysis of variance evaluated differences between sinistrals and dextrals for right and left breasts on pressure-sensitivity; a second analysis of variance analyzed the same variables for two-point discrimination. The first analysis demonstrated a significant difference between the breasts of the combined groups: the left breast was significantly more sensitive than the right breast ($p < 0.001$). There were no differences between the sinistral and dextral groups, nor was the interaction of handedness and laterality significant. Neither handedness, laterality, nor their interaction was significant for two-point discrimination.

Because of the large variability, frequency-comparisons were made, *i.e.* individual *S*s were classified according to whether the right or the left breast was more sensitive for pressure, or for two-point discrimination, a procedure found valuable in a previous study.[12] All 16 of the dextrals and 12 of the 16 sinistrals showed more pressure-sensitivity on the left. A majority of *S*s (10 sinistrals, and 11 dextrals) showed better two-point discrimination on the right. Although neither group differed from chance for this measure, the total of both groups demonstrated a significant majority of *S*s who were more sensitive on the right. Chi-square comparisons of the groups revealed no significant difference for either measure.

Table I shows the patterning of differential sensitivity in terms of the number of *S*s in each group who were more sensitive on the right breast for both measures, and more sensitive on the right breast for one and on the left breast for the other measure. There were no significant deviations from a chance-distribution in the sinistral group, but, in the dextral group, 11 of the 15 *S*s were more sensitive on the left breast for pressure and on the right breast for two-point discrimination ($p < 0.001$). There were no significant differences between the two groups for these patterns.

Two procedures were employed in an attempt to relate total surface-area to sensitivity: Pearson correlations between area and each measure; and χ^2s between groups categorized according to which breast was the larger of the two and which the more sensitive of the two. The first procedure related area to sensitivity within the group; the second evaluated the relationship within the individual.

---

[12] Weinstein and Sersen, *op. cit.*, 665-669.

The correlations were moderately positive, *i.e.* the larger the breast the higher the threshold, but, for the individual *S*, the opposite tendency emerged. Thus, of the 12 dextrals in whom the left breast was the smaller, it was more sensitive for pressure in all, but more sensitive for two-point only in 3.

TABLE I

PATTERNS OF SENSITIVITY OF RIGHT AND LEFT BREASTS FOR TWO MEASURES IN SINISTRALS AND DEXTRALS

R and *L* refer to right and left breasts. The first letter in each pair designates the side more sensitive to pressure; the second designates the side with better two-point discrimination. Thus, *R-L* means the right breast showed better pressure sensitivity and the left better two-point discrimination. One *S* was not tallied because she showed no difference in two-point discrimination.

*S*s	*R-R*	*R-L*	*L-R*	*L-L*
Sinistrals	3	1	7	5
Dextrals	0	0	11	4
Total	3	1	18	9

The tests were repeated in four dextral *S*s. In three *S*s the tests were repeated two more times three weeks to three months later. In another dextral *S*, they were repeated once after three months. In all cases, in each testing, it was found that the left breast was more sensitive than the right breast for pressure-sensitivity and the converse was true for two-point discrimination.

Discussion. The finding of greater sensitivity for pressure on the left of two homologous areas conforms to previous studies.[13] These, however, have dealt exclusively with either highly motile regions (palms and soles), or with other less motile, skeletal regions (forearm and upper arm). The present study demonstrates that the superiority of the left side with regard to cutaneous pressure-sensitivity is not confined solely to these organs.

The converse finding for two-point discrimination has not previously been reported. Of particular interest is the fact that the side which showed the greater sensitivity for pressure was poorer for two-point discrimination. A similar finding was reported by Axelrod in a study dealing with tactual sensitivity of hands of blind and sighted children; the finger which was poorer for two-point discrimination was better for pressure-sensitivity.[14] A more distant analogy is found in vision. Central vision has a greater retinal resolution than peripheral vision but the absolute threshold for perception of light is lower in the peripheral than in the central field.[15] Pressure-sensitivity and two-point discrimination are the tactual equivalents of brightness-perception and visual acuity, respectively. Although the anatomical basis for this difference between peripheral and central fields of vision is known, the anatomical basis for such a sensory difference in the breasts is at present unknown.[16]

---

[13] Semmes, Weinstein, Ghent, and Teuber, *op. cit.*, 6; Ghent, *op. cit.*, 670-673; Weinstein and Sersen, *op. cit.*, 665-669.

Although there were tendencies for the sinistrals to differ from the dextrals in the patterns of sensitivity, these differences were not statistically significant. Weinstein and Sersen found very few differences between sinistrals and dextrals for sensitivity of the hands until the *S*s were differentiated according to their familial backgrounds of sinistrality or dextrality.[17] The number of sinistrals in this study was too few to make an adequate comparison according to familial background possible; when the sinistrals were compared according to familial backgrounds of handedness, there were no apparent differences in their sensory patterns.

When analysis of the sensory patterns of the breasts was made according to the relative cutaneous areas, a paradox emerged. *Within the group,* those *S*s with smaller breasts were more sensitive for two-point as well as for pressure-sensitivity; *within the same S,* the smaller breast was more sensitive for pressure, but the larger breast more sensitive for two-point discrimination. The basis of this discrepancy is unknown. It is thus unclear what role skin-area plays with regard to either of these sensory measures.

SUMMARY

A group of 16 dextral *S*s and a group of 16 sinistral *S*s were studied for pressure-sensitivity and two-point discrimination of the breasts. In a majority of both groups, pressure-discrimination was better on the left and two-point discrimination better on the right. More complex relations between measures of breast-size and sensitivity also were obtained.

---

[14] Seymour Axelrod, *Effects of Early Blindness, Amer. Found. for the Blind Research Series,* No. 7, 1959, 59-61.

[15] S. L. Polyak, *The Retina,* 1941, 221.

[16] M. R. Miller and Michiko Kashahara, The cutaneous innervation of the human female breast, *Anat. Rec.,* 135, 1959, 153-168; personal communication with M. R. Miller reveals that peripheral neuroanatomical differences between right and left breasts were not found.

[17] Weinstein and Sersen, *op. cit.,* 665-669.

CHOICE-BEHAVIOR AND DISTRIBUTION OF REINFORCEMENT AMONG RESPONSE-ALTERNATIVES

By ALLAN R. WAGNER and FRANK A. LOGAN, Yale University

According to the statistical associative model frequently applied to modified verbal-conditioning,[1] each reinforcement of one of the *k*-possible responses results in a specifiable change in the probability of alternative A_i. That is, there is added to the existing probability (P_i) and increment θ $(1 - P_i)$, if A_i is reinforced, or a decrement, $-\theta P_i$, if a response other than A_i is reinforced. This version of statistical associative theory, when applied to a situation in which there are more than two responses, incorporates the simplifying assumption that these changes depend only upon the probability of A_i and are invariant with the distribution of the remaining probability among the alternative responses.

In contrast to this formulation, it may be proposed that in a situation involving more than two possible responses, the size of the decrement in the probability of A_i may depend upon which of the other alternatives is reinforced. Likewise, the incremental effects of a reinforcement of A_i may depend upon the distribution of values of the alternative responses. Such a position is suggested within the context of Hull-Spence theory when it is assumed that the probability of a given response is a function of the probability that the momentary strength of that response is greater than the momentary strength of any one of the alternatives.[2] According to this view, concentrating reinforcements in order to build up a single strong alternative should be more effective in reducing the probability of A_i than distributing those reinforcements among several weaker alternatives.

Two studies which have questioned the adequacy of the statistical associative model might be interpreted as supporting the latter suggestion. Both Gardner, and Cotton and Rechtshoffen have found that the asymptotic frequency of the most often reinforced response is greater if the

---

* Received for publication September 5, 1961. This research was supported by National Science Foundation Grants G13080 and G9014. Our thanks are due Anthony Butterfield, Donald Catlin, and George Farris, who acted as *E*s during the experiment.

[1] See W. K. Estes and J. H. Straughan, Analysis of a verbal conditioning situation in terms of statistical learning theory, *J. exp. Psychol.,* 47, 1954, 225-234.

[2] C. L. Hull, *Principles of Behavior,* 1943, 341.

remaining reinforcements are distributed among two, rather than concentrated on a single alternative.[3] This finding, however, might be attributable to other inherent differences between a two- and a three-choice situation. The following experiment was designed to evaluate the Hullian suggestion when the number of response-alternatives is equated.

METHOD

Apparatus. The apparatus conformed to the general features described by Estes and Straughan.[4] *S* was seated in a sound-proof, dimly illuminated room, facing a circular signal board 12 in. in diameter and centered 4 ft. distant at his eye-level. Around the perimeter of the signal-board were mounted twelve 6.3-v. light-bulbs evenly spaced. A quarter-circular panel placed across the arms of *S*'s chair held five push-button response keys (A_i), and five 6.3-v. lights for 'reinforcement' (E_i). The response-keys were evenly spaced 3 in. apart around the inside of the panel and each of the reinforcing lights was located 2 in. beyond its corresponding key.

Subjects. The *S*s were 24 undergraduate men who served as paid volunteers. Only *S*s with no current or previous enrollment in a course in psychology were used.

Procedure. When S had been seated, he was read the following instructions:

> During the experiment it will be necessary for you to push the buttons in front of you, and to watch both the large board on the wall and small lights by the buttons. The 'ready' signal on each trial will be a flash of the lights on the wall. About a second later, one of the lights on your chair will light up for a moment. As soon as the 'ready' signal flashes, you are to guess which of the lights on your chair will come on, and to indicate your choice by pressing the button corresponding to the light of your choice. If you expect the extreme right lamp to light, press the extreme right button; if you expect the middle lamp to light, press the middle button, and so on. Your object is to guess correctly as many times as you can. It is important that you press one of the buttons on each trial, but never more than one.

After questioning to insure that *S* understood the instructions, he was further told:

> The trials will be run off without any further instructions or conversation. Please make a choice on every trial, even if it seems difficult. Make a guess on the first trial, then try to improve your guesses as you go along and make as many correct choices as possible. There are no tricks or catches in this experiment. We simply want to see how well you can profit from experience in a rather difficult problem-solving situation while working under time pressure.

Following the instructions, all *S*s were given 320 training trials during which only two of the E-lights were used. The *S*s were then randomly assigned to one of two groups of 12 *S*s and with no break in procedure given an additional 160 trials during which all five E-lights appeared. Table I presents the relative frequency, π_i, of each of the E-events during training and during the post-shift condition for the two groups. It should be noted that in both the training and post-shift phases

---

[3] R. A. Gardner, Probability-learning with two and three choices, this JOURNAL, 70, 1957, 174-185; J. W. Cotton and Allan Rechtschoffen, Replication report: Two- and three-choice verbal-conditioning phenomena, *J. exp. Psychol.*, 56, 1958, 96.

[4] Estes and Straughan, *op. cit.*, 225-234.

π_1 and π_5 are the same in Group I as they are in Group II; the groups differ only in the manner in which the remaining post-shift E-events are distributed in π_2, π_3, and π_4. The appropriate relative frequencies of the E-events were constant in each block of 40 trials but within each block the sequence was randomly determined. All *S*s received the same randomized sequence during training. During the post-shift phase the schedules for the two groups were equated with respect to (a) the trials on which E_1 or E_5 occurred, and (b) the trials on which the first E_2, E_3, and E_4 occurred. For half of the *S*s in each group E_1 was on the extreme left, and for half on the extreme right. E_2, E_3, and E_4 always occupied the three middle positions in that order.

The signal, reinforcement, and inter-trial durations were approximately 1, 1, and 2 sec., respectively. The control of these intervals, as well as the programming of the stimuli and the recording of the responses was done manually by *E* in an adjoining room.

TABLE I

SUMMARY OF RELATIVE FREQUENCIES OF OCCURRENCE, (π_i) OF EACH OF THE REINFORCEMENTS (E_i)

Condition	Relative frequency				
	π_1	π_2	π_3	π_4	π_5
Training: All *S*s	.75	.00	.00	.00	.25
Post shift: Group I	.05	.30	.30	.30	.05
Group II	.05	.05	.80	.05	.05

RESULTS AND DISCUSSION

The initial training procedure brought *S*s to a stable, uniform level of responding. There were no reliable changes in the frequency of the several responses over the last six 40-trial blocks of training. In the last such block, the mean frequency of A_1s was 31.7 and 31.8 for Groups I and II, respectively. There were no A_2s, A_3s, or A_4s, leaving 8.3 and 8.2 as the mean frequencies of A_5 for the two groups. As predicted by statistical associative theory, these asymptotic frequencies do not differ significantly from the corresponding frequencies of reinforcement.

Of primary interest was the frequency of A_1 and A_5 responses following the shift in the probabilities of reinforcement. Since there were no differences between the two groups in the pre-shift frequency of these responses or in the post-shift frequencies of reinforcement, π_1 and π_5, the statistical association-model would predict no difference in the post-shift frequency of A_1 or A_5.

The results fail to support this prediction. Table II presents the mean frequencies of occurrence of each of the five responses over all 160 post-shift trials for the two groups. As may be seen the mean frequency of A_1 responses was 16.2 in Group I, but only 6.9 in Group II. Similarly, the mean frequency of A_5 responses was 11.7 in Group I, but only 5.4 in

Group II. Separate Mann-Whitney *U*-tests revealed that the differences between the groups in frequency of A_1 and A_5 responses were significant beyond the 0.1% and 5% levels, respectively.

These results are consistent with the implication of the Hull-Spence formulation. According to this position, the probabilities of A_1 and A_5 would be more greatly reduced in Group II than in Group I following the shift since the Group II values of π_2, π_3, and π_4 would facilitate the acquisition of a response, *i.e.* A_3, which could more effectively compete with the previously preferential responses than any of the several responses in Group I.

TABLE II

MEAN FREQUENCIES OF OCCURRENCE OF EACH OF THE RESPONSES (A_i) OVER ALL 160 POST-SHIFT TRIALS

Group	Frequency				
	A_1	A_2	A_3	A_4	A_5
I	16.2	39.8	48.7	43.6	11.7
II	6.9	7.4	132.3	8.0	5.4

The results of the present investigation are not offered as a critical test of the statistical associative theory *vs.* the Hull-Spence theory. Indeed the latter theory has never seriously been worked out for the experimental situation employed. Presumably a variety of assumptions would be required before the Hull-Spence theory could make predictions concerning such specifics as the final pre-shift probabilities of response which are correctly and exactly predicted by the statistical associative theory. What the data do indicate is that one of the simplifying assumptions frequently made in the statistical associative theory can easily be shown to be inadequate. At the same time, the position apparently necessitated by the post-shift data is basic to the method for deriving probability of a response in the Hull-Spence theory.

SUMMARY

Following identical training in a modified verbal-conditioning situation in which only two responses were reinforced, two groups of *S*s were further trained with five responses reinforced. It was found that the probability of a response depended not only upon the relative frequency of reinforcement of that response, but also upon the distribution of the remaining reinforcements among the alternative responses. This finding was seen to be at variance with a common assumption of statistical association-theory, but consistent with Hull-Spence theory.

DISCRIMINATIVE REACTION-TIME AS A JOINT FUNCTION OF MANIFEST ANXIETY AND INTELLIGENCE

By John R. Stabler, Louisiana State University, and
James A. Dyal, Texas Christian University

In an experiment designed to test the hypothesis that low-drive *S*s should perform better than high-drive *S*s in complex tasks, Grice found an inverse relation between discriminative reaction-time (*DRT*) and anxiety-level, as measured by Taylor's Manifest Anxiety Scale (*MAS*).[1] When the data were co-varied upon an intelligence-measure, however, the relationship disappeared.

One implication of Grice's work was that the positive results of other tests of the same hypothesis might be accounted for in terms of intelligence rather than motivation. The purpose of the present study was to test this hypothesis under conditions in which high- and low- anxiety groups were matched for intelligence.

Method. (1) Subjects. The *S*s were men, all prisoners at Seagoville Federal Prison in Seagoville, Texas. Of 265 *S*s available for *MAS*-testing, 36 were eliminated because of high scores on the *L* and *K* scales, and 33 because they either misunderstood directions or refused to participate; thus, 196 *S*s satisfactorily completed the *MAS*. Four experimental groups of 15 *S*s each were selected from this sample. The mean *MAS*- and *IQ*-scores of the respective groups are presented in Table I. The *S*s responded to items of the *MAS* by placing check marks on 7-point scales which rep-

TABLE I

Mean *MAS*- and *IQ*-Scores

Group	*MAS*		*IQ*	
	Mean	*SD*	Mean	*SD*
High *MAS*, High *IQ*	+37	15.8	108	8.3
High *MAS*, Low *IQ*	+37	13.9	84	6.3
Low *MAS*, High *IQ*	−86	18.7	112	6.4
Low *MAS*, Low *IQ*	−84	18.0	84	6.1

---

* Received for publication October 2, 1961. The coöperation of the staff and prisoners at the Seagoville Federal Prison, Seagoville, Texas, is gratefully acknowledged.

[1] G. R. Grice, Discrimination reaction time as a function of anxiety and intelligence, *J. abnor. soc. Psychol.*, 50, 1955, 71-74; J. A. Taylor, A personality scale of manifest anxiety, *J. abnorm. soc. Psychol.*, 48, 1953, 285-290.

resented the degree of truth or falsity of the *MAS* items. A check at the midpoint of a scale was scored zero; while other check marks were scored either 1, 2, or 3, depending upon their displacement from the midpoint. The sign given the score, *plus* or *minus,* depended upon whether or not agreement with the item indicated presence of anxiety. The scores of two *S*s, subsequently omitted because of illness and uncoöperativeness, were divided proportionately among the *S*s remaining within their respective groups.

Apparatus. The apparatus consisted of a stimulus-display, a response-panel, and a Standard Electric timer. The stimulus-display consisted of three colored lights arranged vertically in the center of a dull-black plywood panel. They were separated from one another by 2 in. and from the response-panel below by 12 in. A white warning light was located 6 in. to the left of the stimulus-lights. The response-panel consisted of three buttons in a triangular formation and separated from each other by 1 in. Two short strips of masking tape arranged to form an 'X' were fixed to the table 6 in. in front of the response-panel.

Procedure. A white warning light was presented before the stimulus-light. The interval between the warning and stimulus-lights was varied randomly in 0.5-sec. steps from 0.5 to 3.5 sec. *S* responded to the stimulus-light by moving his hand from the masking tape to the appropriate response button as quickly as he could. Pressing the correct button turned off the stimulus-light. The order of presentation of stimulus-lights was random within each block of 15 trials except for the one restriction that all lights appeared an equal number of times. *S* rested his hand on the masking tape between trials. Discriminative reaction-time and errors were recorded after each trial.

The testing session was conducted in two stages. Each *S* received 90 trials with one combination of *S–R* pairs during Stage 1 and 60 trials with another combination of *S–R* pairs during Stage 2. The second stage came immediately after the first.

Results. The performance of the four groups during the first and second stages is shown in Fig. 1 in terms of mean speed of response ($1/DRT$) for 15-trial blocks. Analysis of variance of the data of Stage 1 showed reliable *IQ*-effects: ($F = 5.57, p < 0.05$), Trial- ($F = 107.57, p < 0.001$), and *MAS* × Trial- ($F = 3.67, p < 0.01$).

Although the graph in Fig. 1 indicates that *DRT*-performance of low-anxiety *S*s surpassed that of high-anxiety *S*s at both *IQ*-levels, a Duncan range-test showed that the low-anxiety *S*s were reliably faster only on the first three and fifth blocks of trials.[2] The Duncan range-test also showed that high-anxiety *S*s improved their speeds significantly on each block of trials up to the fifth block, while low-anxiety *S*s improved significantly only from Block 1 to Block 2.

To determine whether or not the change from Stage 1 to Stage 2 differentially affected the anxiety-groups, an analysis of variance was performed which included only the last 15 trials of Stage 1 and the first 15 trials of

---

[2] D. B. Duncan, Multiple range and multiple *F*-test, *Biometrics,* 11, 1955, 1-42.

Stage 2, *i.e.* Blocks 6 and 7. Reliable *IQ*- ($F = 7.98, p < 0.01$) and Trial-effects ($F = 127.30$, $p < 0.001$) were obtained, but the interaction was insignificant. Analyses of mean response-speeds during Stage 2 also showed reliable *IQ*- ($F = 7.62$, $p < 0.01$) and Trial-effects ($F = 42.8$, $p < 0.001$). Statistically significant Trial-effects, but not reliable *IQ*- or *MAS*-effects, were obtained from analyses of the reciprocal mean frequency of errors in Stage 1 and Stage 2 ($F = 58.80$, $p < 0.001$, and $F = 74.33$, $p < 0.001$, respectively).

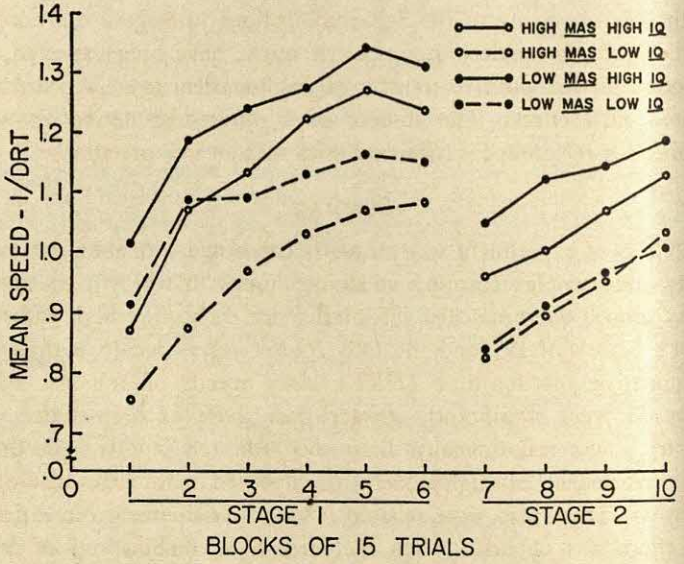


FIG. 1. MEAN SPEEDS OF THE FOUR EXPERIMENTAL GROUPS

The correlation between *IQ*- and *MAS*-scores, based upon the original 196 *S*s, was not statistically significant ($r = -0.06$).

Discussion. The inverse relationship between the discriminative reaction-time (DRT) and anxiety-level found on the first three blocks of trials supports the hypothesis of an interaction between motivation and complex learning. The loss of a reliable *MAS*-effect on later trials suggests that the advantage of low-anxiety *S*s diminished as the task became easier through repeated practice. As correct S–R tendencies strengthened, indicated by increasing response-speeds and decreasing error-frequencies, the *DRT*-performance of high-anxiety *S*s improved more than that of low-anxiety *S*s.

The Yerkes-Dodson law holds that optimal motivation for a learning task decreases with increasing difficulty.[3] The present findings for Stage 1 are in accord with the law in that low-anxiety *S*s performed better than high-anxiety *S*s when task-difficulty was maximal; as the task became easier through repeated experience, and high-anxiety *S*s improved more than low-anxiety *S*s, the advantage of low motivation was lost.

Although re-pairing of the S–R combinations in Stage 2 did not differentially affect the anxiety groups, as it might have been expected to do, it is possible that positive transfer of habit-tendencies of a board nature obscured such effects. The absence of a differential decrement in performance for *IQ*-groups is consistent with such an interpretation.

SUMMARY

The present experiment was primarily concerned with the interaction of anxiety and complex learning under conditions in which high- and low-anxiety groups were matched for intelligence. *S*s scoring high and low on Taylor's Scale (*MAS*) and the Otis *IQ*-test responded to a three-choice discriminative reaction-time (*DRT*) task. Speeds of response of low-anxiety *S*s were significantly greater than those of high-anxiety *S*s on early trials. As task-difficulty diminished through repeated practice, the *DRT*-performance of high-anxiety *S*s improved more than that of low-anxiety *S*s. There also were reliable *IQ*- and Trial-effects. No differential *MAS*-effect was obtained when the three S–R combinations of Stage 1 were re-paired to form new combinations for Stage 2. The results were interpreted in terms of the Yerkes-Dodson law and changes in task-difficulty as a function of repeated practice.

---

[3] R. M. Yerkes and J. D. Dodson, The relation of strength of stimulus to rapidity of habit-formation, *J. comp. neurol. Psychol.*, 18, 1908, 459-482.

RECOGNITION OF TACTUAL FORM BY SIGHTED AND BLIND SUBJECTS

By ANNE G. EWART and FRANCES M. CARP, Trinity University

Compensatory development of the other senses—especially tactual and proprioceptive—in the blind has long been a subject of investigation.

In regard to tactual recognition of form, comparisons of the performance of blind and sighted subjects (*Ss*) are conflicting. Plata, using the two-point limen and difference-thresholds for lifted-weights and estimated heights, concluded that blind *Ss* were inferior to sighted.[1] On the other hand, Brown and Stratton found some advantage for the blind in regard to the two-point limen.[2] Bartley *et al.* reported no difference in judging the size of blocks.[3] Hunter's blind *Ss* were more accurate in perception of a 'plus curve.'[4] Worchel found no difference in perception of simple geometric forms.[5]

This study attempts to compare tactual recognition of form in the sighted and the blind, avoiding some of the difficulties in previous studies. Bartley used one small sample (10 blind and 10 sighted), Plata's controls were inadequate; Hunter used an inflated *N* (number of judgments rather than of *Ss*) in statistical analysis, and Worchel's tasks did not differentiate the preferred scores of his *Ss*.

Subjects. The *Ss* were 30 children in the Texas State School for the Blind and, to control for institutionalization, 30 sighted children in a children's home. Blindness was defined as total visual loss or light-perception only. Twenty-four *Ss* were congenitally blind.

The mean chronological age (*CA*) of the blind *Ss* was 11.9 yr. (*SD*, 2.4 yr.; range, 8 to 16 yr.), and their mean *IQ* was 111 (*S.D.*, 18.1; range, 80 to 145); mean *CA* of the sighted *Ss* was 11 yr. (*SD*, 2.0 yr.; range 8 to 16 yr.), and mean *IQ*, 101, (*SD*, 14.5; range, 81 to 140). No attempt was made to equate groups for sex, since previous studies report no differences. There were 13 boys and 17 girls in the blind, and 18 boys and 12 girls in the sighted group.

The *IQ*s of sighted *Ss* were measured by the Stanford-Binet Scale, Form M; that of blind *Ss* by either the Interim Hayes-Binet or the WISC (Verbal) Scale. In the absence of an instrument for equating *IQ*s of the blind and sighted, intelligence was controlled only to the extent that the lower levels (below 80) were elimi-

---

* Received for publication December 11, 1961.

[1] C. E. Seashore and T. L. Linn, The comparative sensitiveness of blind and seeing persons, *Psychol. Monogr.*, 25, 1918, (No. 108), 148-158.

[2] M. S. Brown and G. M. Stratton, The spatial threshold of touch in blind and seeing children, *J. exp. Psychol.*, 8, 1925, 434-442.

[3] S. H. Bartley, The perception of size or distance based on tactile and kinesthetic data, *J. Psychol.*, 36, 1953, 401-408; S. H. Bartley, L. T. Clifford, and A. D. Calvin, Effect of visual imagery on tactual and kinaesthetic space perception, *Percept. Mot. Skills*, 5, 1955, 177-184.

[4] I. M. L. Hunter, Tactile-kinaesthetic perception of straightness in blind and sighted humans, *Quart. J. exp. Psychol.*, 6, 1954, 149-154.

[5] Philip Worchel and Jack Mauney, The effect of practice on the perception of obstacles by the blind, *J. exp. Psychol.*, 41, 1951, 170-176.

nated. It is evident that the mean *IQ* for the blind group is higher and the distribution of their scores less skewed. The effect of possible differences in intelligence between groups is considered in the interpretation of results.

Stimulus-material. Eight stimulus-forms were used: circle (1-in. radius), square (2 × 2 in.), semicircle (2-in. radius), rectangle (1.5 × 2.5 in.), quarter-circle (2-in. radius), equilateral triangle (2-in. sides), crescent (2 in. between tips) and parallelogram (2-in. sides). There were four choice-blocks for each stimulus-block, one being identical in shape with the stimulus-form. Choice-blocks cut from ½-in. hardwood, sanded and varnished to a smooth finish, were approximately twice the size of the stimulus-block. Differences between correct and incorrect choice-forms varied, but were very slight for at least two. Final design of the choice-forms was determined by pre-tests with sighted children not used as *S*s.

Procedure. The general purpose of the test was explained individually to *S*. Then he was blindfolded (blind and sighted alike) and given the following instructions:

> *Instructions.* I am going to hand you a block. I want you to feel it so you can remember the shape. You may use one or both hands. When you lay the block down, I will hand you four blocks, one at a time. One of these four blocks has the same shape as the one you just felt. I wish you to pick out the one of these four blocks that has the same shape as the first one and hand it to me. You may feel them as often as you like until you are sure that you have the one that has the same shape. Do you understand?

S was then given a stimulus-block. When he returned it, the four appropriate choice-blocks were given him. The position of the correct block within the series of presentations was systematically varied. *S* was permitted to handle the choice-blocks as long as he desired.

Results. The data first were treated by analysis of variance, categorized in terms of Vision (blind vs. sighted) and *IQ* (high vs. low). If an *S*'s *IQ* was above the median *for his own group,* it was considered high; if below, it was regarded as low. This *IQ*-categorization was based on the temporary assumption that the groups were comparable in intelligence and that the differences in *IQ* were due to the tests and administrative procedures. For the Vision × *IQ* categorization, proportionality was achieved and the assumption of homogeniety of variance could not be rejected at the 2% level of confidence. For testing main effects and interactions, the 5% level of confidence was selected.

The variances attributable to Vision and *IQ* were well below the established level of statistical significance (respectively, $F = 2.77$, $df = 1$; $F = 2.25$, $df = 1$). Performance was not significantly affected by intelligence-score when blind and sighted were considered together, and there was no difference between the blind and the sighted. The interaction-variance, Vision × *IQ,* however, was significant at the 5% level ($F = 4.00$, $df = 1$). The high *IQ* blind *S*s (mean score 5.40) were superior not only to the low *IQ* blind (mean score 4.00), but to both the low *IQ* (mean score 4.13) and the high *IQ* (mean score 3.93) sighted *S*s. These differences, using the *t*-test with the error-term from the analysis of variance, were significant at better than the 1% level of confidence.

The data also were treated by Pearsonian correlation-technique, which would be less affected by the disparity in *IQ* between the two samples than would analysis of variance. *CA* was included as a variable in these analyses. Results corroborated those of the analysis of variance. The coefficients for *CA* vs. form-recognition were identical and low for both groups ($r = 0.24$). The correlation between *IQ* and *CA* was low negative for both groups ($r = -0.31$, $r = -0.13$). None of these correlations is significant at the 5% level. As suggested by the analysis of variance, however, the correlation between *IQ* and form-recognition was sizable for the blind *S*s ($r = 0.54$) but negligible ($r = 0.01$) for the sighted. The difference between these coefficients, after transformation to Fisher's *z,* is significant at better than the 5% level of confidence.

The partial correlation between *CA* and form-recognition, with *IQ* held constant, was 0.26 ($P > 0.05$) for the sighted, and 0.37 ($P > 0.05$) for the blind. The partial correlation between *IQ* and form-recognition, with age held constant, was 0.09 ($P > 0.05$) for the sighted sample, but 0.59 ($P < 0.01$) for the blind. This further supports the analysis of variance.

Discussion. Using materials sufficiently difficult to obtain a range of scores, this study confirms Worchel's conclusion of no difference between blind and sighted children in tactual form-perception (using the method of recognition). It supports Worchel's findings, rather than those of Plata, Hunter, and others, in showing insignificant correlations between chronological age and score.

Significant interaction was demonstrated between *IQ* and vision. High *IQ* blind *S*s were superior in tactual recognition of form, not only to low *IQ* blind, but to all sighted *S*s. For the blind, there was a significant relationship between *IQ* and tactual recognition of form; among the sighted there was not. Comparing the 15 blind and 7 sighted children with *IQ* scores of 110 or more, the mean tactual recognition-score of the sighted *S*s was 4.2 (differing only slightly from the mean score of the entire high *IQ* sighted group) while that of the blind was 5.4.

Both blind and sighted *S*s made the most errors on the semicircle, quarter-circle, and triangle, the fewest on the circle and crescent. With these forms, mean errors for the blind and sighted were essentially the same.

The forms which accounted for most of the differential variance between blind and sighted *S*s were the square, rectangle, and parallelogram. On the basis of visual inspection, the differences between the one correct and three incorrect forms seem apparent. Yet when blindfolded and dependent on tactual cues, sighted *S*s had considerable difficulty in selecting the correct form. This may be due to some fortuitous circumstance in the construction of the three series of rectangular forms. On the other hand, it may result from an experience-difference of blind and sighted *S*s.

For sighted *S*s, rectangles are predominant in everyday experience. Doors and windows, buildings, and streets are essentially rectangular. Only from one position is an object truly rectangular, but it is perceived visually as such from other angles. For sighted *S*s this tendency to categorize as rectangular objects which deviate somewhat from rectangularity may generalize to tactual perception. Form-concepts in the blind, based on tactual and kinesthetic experience, are not so likely to provide for categorizing objects as rectangular on the basis of sensory impressions of non-rectangularity. Consequently, these form-concepts should be more distinct for the blind than for the sighted and categorization of percepts more accurate. Among blind *S*s, the more intelligent apparently had developed more distinct and complete form-concepts and greater facility in relating new stimulus patterns to them.

SUMMARY

Tactual recognition of form was tested with 30 blind and 30 sighted children, using a series of 8 stimulus-blocks, each of which was accompanied by four choice-forms, one of which was identical in shape to the stimulus-form.

The results of this study confirm Worchel's conclusion that visual imagery is not a critical factor in this type of form-recognition. Analysis of variance and correlational analysis showed no significant effect of vision, *IQ*, or *CA* on the test-scores. The interaction between vision and *IQ* was, however, significant. Though there was no difference in scores between sighted *S*s with high and low *IQ*s; blind *S*s with high *IQ*s were superior not only to the blind with low *IQ*s, but also to the sighted *S*s—both high and low *IQ*s. Differences in shape of forms which were most readily recognized by the blind *S*s as compared to the sighted *S*s suggest that some form-concepts developed largely on the basis of visual percepts may be over-inclusive for tactual purposes.

THE INFLUENCE OF COLOR ON DISCRIMINATION OF SWEETNESS IN DRY TABLE-WINE

By Rose M. Pangborn, Harold W. Berg, and Brenda Hansen,
University of California, Davis

Previous investigations showed that the ability to discriminate between sugar-levels in aqueous solutions was not influenced by red, green, or yellow food-coloring.[1] In pear-nectar, one group of *O*s considered the green-colored samples as least sweet and the colorless nectar as sweetest,[2] whereas a second group of *O*s demonstrated neither color-sweetness nor color-sourness associations in pear-nectar.[3] In this latter study there was a pronounced trend toward less accuracy of identification of taste in colored than uncolored nectars indicating that visual variables interfered with gustatory responses.

The present paper reports the effect on discrimination of sweetness of coloring dry, white table-wine. To simulate sauterne, sherry, rosé, claret, and burgundy, the white wine was colored yellow, brown, pink, red, and purple, respectively.

Selection of panel and training. The responses of two separate panels were obtained: (a) eight men and three women, highly experienced in taste-testing of non-alcoholic beverages but naïve to judging wine, hereafter referred to as naïve wine-tasters; (b) four men with several years' experience in the evaluation of the tastes of foreign and domestic wines, hereafter called the experienced wine-tasters. With both panels, the first five days were devoted to training, *i.e.* to acquainting the *O*s with the judging system and to accustom them to recognize small differences in sweetness in uncolored samples of wine. Judging was done daily, Monday through Friday at 3 P.M. in individual, temperature-controlled, partitioned booths. Each booth was illuminated with two 100-w. incandescent bulbs mounted above a translucent glass-plate which served to diffuse the light. At each session, the *O*s received four randomly presented, paired sets and were instructed to circle the number of the sweeter sample within each pair. Wines were stored at refrigeration-temperatures, but allowed to reach room-temperature (approx. 74°F.) before tasting. Distilled water at room-temperature was provided for oral rinsing. The *O*s were requested to retaste the wines as often as they wished but to refrain from swallowing. Im-

---

* Received for publication January 9, 1962.

[1] R. M. Pangborn, Influence of color on discrimination of sweetness, this Journal, 73, 1960, 229-238.

[2] *Ibid.*, 237.

[3] R. M. Pangborn and Brenda Hansen, The influence of color on discrimination of sweetness and sourness in pear-nectar, this Journal, 76, 1963, 315-317.

mediately after each session of tasting, *O*s were informed as to the accuracy of their responses.

Preparation of samples. The dry, white table-wine used throughout these experiments was purchased from a commercial California winery and had the following composition: 0.51% total acid (calculated as tartaric acid), 12.2% alcohol, 1.7% extract, 41 ppm free SO_2, and less than 0.2% residual sugar. Table I lists the codes and amounts of sucrose and food-coloring used to prepare the yellow, brown, pink, red, and purple colors which were indistinguishable visually from sauterne, sherry, rosé, claret, and burgundy wine-types, respectively. To check the accuracy of the preparation of the samples, each batch was tested separately for sugar-content by refractometry and for color by colorimetry.

TABLE I

STIMULUS-SOLUTIONS (*A–L*) AND THEIR COMPOSITION

Color composition		Amt. base added	% Sucrose†	
Hue	Base*	to 1 L. wine	1.5	1.8
White	—	—	A	G
Yellow	2 ml. orange 0.32 ml. blue 0.24 ml. red	10 ml.	B	H
Brown	3.2 ml. orange 0.4 ml. blue 1.0 ml. red	12.5 ml.	C	I
Pink	0.2 ml. blue 4.0 ml. red	11 ml.	D	J
Red	0.2 ml. blue 8.0 ml. red	45 ml.	E	K
Purple	4.0 ml. blue 20.0 ml. red	50 ml.	F	L

* Amount of Schillings food-coloring added to 100 ml. of wine.
† Study I: 1.5 vs. 1.8% sucrose, naïve and experienced *O*s.
Study II: 5.0 vs. 5.4% sucrose, naïve *O*s only.
Study III: 2.0 vs. 2.0% sucrose, naïve and experienced *O*s.

As indicated in the footnote in Table I, sucrose-levels of 1.5 vs. 1.8% were evaluated by both the naïve and experienced *O*s; sucrose-levels of 5.0 vs. 5.4% were evaluated by the naïve *O*s only; and samples identical in sucrose-content (2.0%) but differing in color were judged by both groups.

The 36 paired combinations were evaluated in sets of 4. Nine days were required for each judge to evaluate each set. Chi-square analysis was applied to the data to determine whether observed values differed significantly from expected values.

Taste-discrimination in colored and uncolored wines. Reactions to sweetness in colored and uncolored wines are given in Table II. Among the naïve *O*s, three men had a much higher accuracy of selection of the sweeter samples in uncolored than in colored wines when the sucrose-levels were 1.5 vs. 1.8%. When the sucrose-content was 5.0 vs. 5.4%, only one *O* exhibited a greater accuracy of selection within uncolored pairs. For two *O*s, a reversal of the distribution mentioned above was recorded The experienced *O*s performed better when evaluating sweetness in colored wines, but this may have been an effect of learning. In general, no specific pattern of average response was demonstrated. Values of correct responses ranged from 43 to 87% for the naïve panel and from 66 to 90% for the experienced *O*s.

TABLE II

RESPONSE TO SWEETNESS IN COLORED AND UNCOLORED WINE

*O*s	% Sucrose	Uncolored	Colored	*t*
Naïve	1.5 vs. 1.8	60.4±15.6*	58.6±16.3*	0.252
	5.0 vs. 5.4	60.6±16.3*	60.0±13.8*	0.031
Experienced	1.5 vs. 1.8	70.2±19.6†	78.8±10.4†	0.853

* Significant at the 5% level.
† Significant at the 0.1% level.

TABLE III

EFFECT OF COLOR ON *O*'S RESPONSE TO SWEETNESS
(Values in %, 16.6% expected.)

	Set I (1.5 vs. 1.8% sucrose)			Set II (2.0 vs. 2.0% sucrose)			Set III (5.0 vs. 5.4% sucrose)
Color	Naïve *O*s (60 trials)	Exp. *O*s (30 trials)	Total (90 trials)	Naïve *O*s (66 trials)	Exp. *O*s (30 trials)	Total (96 trials)	Naïve *O*s (66 trials)
White	19.7	16.1	18.6	14.1	18.9	15.6	18.9
Yellow	17.7	16.7	17.4	15.1	15.55	15.3	14.9
Brown	16.7	17.2	16.8	17.2	15.55	16.7	16.9
Pink	14.9	17.2	15.6	18.2	22.8	19.6	17.5
Red	15.7	16.1	15.8	19.7	12.8	17.5	15.9
Purple	15.3	16.7	15.8	15.7	14.4	15.3	15.9

Color-sweetness associations. The data presented in Table III are expressed as percentage of responses to the intensity of sweetness within each color of wine. Analyis by Chi-square was applied to the individual numbers. The panels responded differentially to the colored wines, but none of the observed distributions, individually or collectively, was significantly different from the expected distributions. One naïve *O* consistently ascribed the greatest sweetness to white, and the least sweetness to red-colored, wines. The Chi-square values for the experienced panel were lower than those calculated for the naïve group indicating that the former group was influenced more by the color variables.

In Sets I and II, similarities in distribution of responses are seen only in that the white wines received the highest percentages of total response. The combined distribution of both panels revealed that white wines were considered sweetest in Set I and pink-colored wine in Set II. Experienced *O*s ascribed the greatest sweetness to the pink wine, possibly through their familiarity with sweet rosé wines. A significant over-selection of the second samply within a pair was demonstrated by two *O*s, while over-selection of the first sample was demonstrated by one *O*. The totals for the panels showed no significant time-error.

APPARATUS

A COLOR-MIXER WITH MONOCHROMATIC PRIMARIES

By JOHN KRAUSKOPF, Rutgers University

The facts of color-mixture can be accounted for, at least to a good first order approximation, by the laws of color-mixture in the *CIE*-system. These relationships constitute a beautiful example of lawfulness in the psychological realm, yet students in undergraduate laboratories often fail to appreciate the significance of the experiment on color-mixture. To a large degree, this failure may be attributed to the lack of vividness of the demonstrations available with, and to the cumbersomeness of, the apparatus employed. The present apparatus remedies these defects.

The principal problem solved in the development of the present apparatus was that of obtaining nearly monochromatic light at a modest cost. Of the methods of producing continuously variable monochromatic light, the use of diffraction gratings recommends itself highly. The spectrum produced is linearly arrayed with respect to wave-length, which means wave-length and purity can be varied easily by linear-metric controls. Relatively high efficiencies are obtained. Stray radiations (at undesired wave-lengths) are relatively low in level. The principal reason that gratings have not been used widely is that they are expensive. Only the fact that plastic replicas of good quality can be made from master gratings at greatly reduced cost makes possible the routine use of diffraction gratings in research monochromators, but the cost of these materials is still outside the limits of many laboratories. Fortunately, lower grade of replica recently has become available at extremely low cost.[1] While these gratings are not of research-quality, they have been found to be satisfactory for the present purposes.

Description of existing model. The use of the gratings in the production of monochromatic light is illustrated in Fig. 1. Light from a lamp (S) with a vertical ribbon filament (GE 18A T10 Fil.) passes through a shutter and is collimated by lens L_1. It then passes through the diffraction grating (G)—a 2 × 2 in. piece of

---

* The development of this apparatus was supported in part by the U.S. Army Medical Research and Development Command, Department of the Army, under Research Contract No. Da-49-193-MD-2128. The author is now with the Division of Neuropsychiatry, Walter Reed Army Institute of Research, Washington, D.C.

[1] These gratings may be purchased from the Edmund Scientific Co., Barrington, N.J.

grating mounted in a glass sandwich—and then through Lens L_2. Spectra of varying order are laid out in the focal plane of L_2. The brightest (zero-order) spectrum appears directly on the line from S through L_1 but is not useful since it is white. First-order, second-order, and higher-order spectra, in which the light is dispersed with respect to wave-length, are arrayed to the left and right of this line. By virtue of the orientation of G and L_2, the first-order spectrum to the right is the most useful. This spectrum is laid out in the plane marked SP. With the optics used ($F_{L1} = F_{L2} = 25$ cm.), the spectrum is approximately 50 mm. wide and its height is equal to the length of the lamp-filament—about 20 mm. A slot-plate is located in the plane of S. A front view of this plate is shown in the insert of Fig. 1. Three slots were milled in this plate. They were so located that one slot passed light from the blue region of the spectrum, one from the green, and one from the red. The amount of light in each of these primaries is varied by occluding the slots with brass strips which slide up and down in grooves in the slot-plate. The strips move freely in the grooves and are suspended by high-test fish-line which in turn is attached to, and wrapped around, the shafts of ¼-in. drill-rods. These shafts extend out *O*'s end of the box and terminate in knobs. The intensity of the three primaries may therefore be controlled by *O*, and the settings of the knobs read to indicate the amount of each primary in instrument-units.

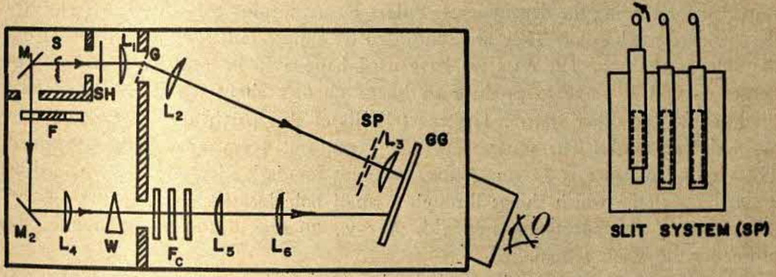


FIG. 1. DIAGRAM OF THE APPARATUS

The light passing through the slots is gathered by L_3 and three nearly superimposed images of the grating in red, green, and blue are formed in the plane of a ground-glass (GG). By suitable masking with black tape, a homogeneous square area is defined on the ground-glass. Two minor defects are evident when this square is viewed from the *O*'s position. The appearance of the square is not perfectly homogeneous, *i.e.* vertical striations are seen which are not all the same color. This effect, which seems to be due to grating-imperfections, can be reduced greatly by using two thicknesses of ground-glass and separating them with a short air-space. The increased diffusion washes out the inhomogeneities rather well. The second defect is a vignetting effect. At the very lowest levels, a shadow-image of the edge of the brass strip is seen in the square; that is, the top of the square brightens before the bottom. This defect, which is noticeable only at low light-levels, could be eliminated by replacing L_3 with two lenses, one in front and one behind SP such that the combination has about the same focal length of L_3 but with its principal plane located in SP.

A second square of light, slightly displaced from the first on GG, is provided by the remaining optical devices. This field is used as the standard field in color-matching and as a variable brightness test-patch in experiments on dark-adaptation. Light for this patch is derived from the back of the lamp, reflected by mirror (M_1) through a neutral filter wheel (F). Homogeneously exposed negative-material was used to provide neutral-density filters. This material was used to cover holes located in the periphery of a masonite disk. Steps in density were obtained by using varying numbers of negatives over each hole. The disk was mounted on a shaft which extended out of the box. The density-level may be controlled by rotation of this shaft. After passing through the filter, the light is reflected toward L_4 which is located at *S*'s focal distance. The collimated beam thus produced then passes through a circular neutral-density wedge which provides continuous control of brightness. (The wedges were made of negative material and photographically copied from research-grade glass wedges.) This wedge is mounted on a shaft which runs the length of the box and may be controlled either by *E* or by *O*. At *E*'s end, a dial-plate and pointer are provided for reading the density-level. Filters F_c are surplus color-filters used to vary the standard-patch color. They are controlled by simple pull-rods which extend out the side of the box. The ones we have used happen to be red, blue, and yellow. Lenses L_5 and L_6 serve to produce an image on GG about the same size as that produced by the other system. The exact limits of this patch are again determined by black-tape masks. The squares are 1 × 1 cm. and 4 cm. apart, center-to-center. The viewing distance is 25 cm. A fixation-point for dark-adaptation is provided by a small 6-v. lamp which shines through a small hole located 8 cm. to the right of the center of the adaptation test-field. A 6-v., 20-amp. transformer supplies the power for the whole system.

When the apparatus is used as a dark-adaptometer, the grating system is shut off. This is accomplished by closing the shutter SH. The shaded area represents a light-baffle which has two exit-holes. The shutter consists of a piece of masonite mounted on a shaft extending out of the box which can be rotated to cover the hole leading to the grating system.

Experiments. Dark-adaptation experiments have been conducted with the apparatus in traditional ways. *O* is light-adapted and then makes repeated ascending adjustments of the circular wedge by rotating the knob protruding from his end of the apparatus. *E* records these settings and makes appropriate level-adjustments of F from time to time. The experiments can be performed with white or colored light by appropriate selection of F_c.

As the apparatus now is arranged, simple color-matching can be achieved, using white or colored light provided by the channel of the adaptometer. Once matches to white are achieved, the complementaries with respect to white may be demonstrated by positive and negative rotation of each of the primary controls. After-image complementaries may be demonstrated by presenting each of the primaries alone and asking *O* to fixate first the primary and then the white square.

The apparatus might serve to demonstrate the effects of light-intensity and color on acuity, but it has not yet been put to this use.

Defects and proposed modifications. The primaries obtained turned out to be too good with respect to the filters used in the matching-channel. It was hoped to be able to demonstrate the need for 'negative' mixture, but all the samples could be matched by positive methods. Further, it is evident that negative mixtures could not have been achieved with the present arrrangement.

One basic change in the apparatus which would greatly enhance its value is the addition in the standard channel of a monochromatic beam of variable wave-length. This could be achieved quite simply by using a configuration similar to that in the primary system. The principal difference would be that one slot, variable in horizontal position, would replace the three fixed slots. Color-matching experiments against this source would provide direct experience with the fundamental measurements that underlie the *CIE*-system, since, in general, matches could not be obtained with the positive mixtures of the primaries. Desaturation with the already extant white or filtered beam would be required, followed by matching of the desaturating beam alone, and a subtraction of these primary values from the original match.

A second way to produce a three-channel instrument might be to take advantage of the fact mentioned above that several orders of spectra are produced by a diffraction grating. The zero-order (white) spectrum is formed on the optic axis with first-, second-, and higher-order spectra arrayed both to the left and right of the axis. One of the first-order spectra might be used to obtain the primaries as in the present apparatus. The other first-order spectrum might be used as a source of selected bands of monochromatic light to be matched. The zero-order spectrum might be joined with this by a suitable arrangement of mirrors and beam-splitters. The zero-order beam could then be used for desaturation of the spectral light or as a source of white light for adaptation-experiments.

A MODIFIED WISCONSIN GENERAL TEST APPARATUS FOR THE RAT

By A. Robert Rollin, Bryan E. Shepp, and K. E. Thaller,
University of Connecticut

The Wisconsin General Test Apparatus (WGTA) has, for many years, been a standard apparatus for studies of discriminative learning and related processes in primates.[1] We describe here a modified version of the WGTA which permits analogous work on object-discrimination in rats. An isometric view of the apparatus is presented in Fig. 1.

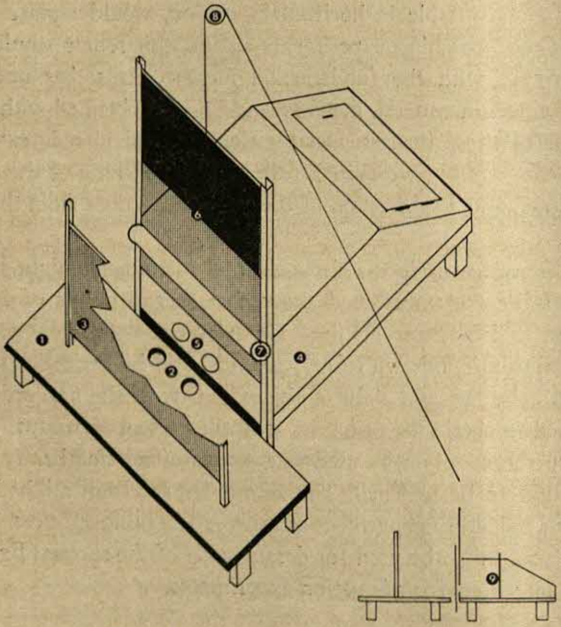


Fig. 1. Modified WGTA for the Rat
(For description, see text.)

Components. The stimulus-table (1)—numbers identify the components in Fig. 1—is a 30 × 12 × 8-in. wooden structure which contains two food-wells (2) and a one-way-vision screen (3). The leading edge of the tray stands juxtaposed to the front of the housing-unit (4). The food-wells are 3/8 in. deep, 7/8 in. in diameter, and separated 2½ in. from center to center. The one-way screen is manually raised and lowered along 3/8-in. aluminum runners, and the screen is held at any position by counterbalancing weights (not shown). The 1/8-in. Plexiglas screen is covered with half-silvered paper to give the one-way vision effect. The table is painted flat white.

---

[1] H. F. Harlow, The formation of learning sets, *Psychol. Rev.,* 56, 1949, 51-65.

The housing-unit (4) is a modified laboratory cage, 9 × 9 × 15 in. The front wall of the unit is constructed of 1/8-in. Plexiglas, while the remaining walls are covered by 1/8-in. masonite. The floor—which is suspended over a litter-pan—and the ceiling are ¼-in. wire mesh. A manually operated door (9)—in side view of Fig. 1—confines the rat to a space, 9 × 9 × 7 in., in the front of the unit. All surfaces excepting the Plexiglas front are painted flat black.

Two windows (5) cut into the Plexiglas front allow the rat to dislodge the stimulus-objects from the food-wells and obtain a reward-pellet. Each window is directly in line with a food-well. The windows are so constructed (1¼ in. in diameter) as to permit the rat to extend its head but to prevent escape to the stimulus-table. (This size of window is suitable for male rats aged 90-150 days; adjustment may be required for older or younger rats.) The distance between windows prevents the rat from dislodging a stimulus-object from either food-well from the opposite window. The bottom of each window is ½ in. above the stimulus-table. A 6-w. standard cool-white fluorescent lamp (7) is suspended over the windows to provide illumination. The windows are covered between trials by a Plexiglas door (6) which slides along aluminum runners attached to the housing-unit. Since the upper portion of the door is painted flat white, the rat cannot see the stimulus-table or its contents when the door is closed. The lower portion of the door is transparent, permitting the rat to see the stimulus-objects before it can respond to either of them. The door is raised at a constant rate (1 in./sec.) by an electric motor (not shown) which operates a string-and-pulley (8). The motor runs at a constant speed and is capable of instantaneous start and stop. The time required to raise the door is 4 sec., and it travels a distance of 4 in.

Operation. The adaptation and pretraining procedures for training rats to perform in the modified WGTA are similar to those procedures employed with the standard WGTA. After several days on a 24-hr. feeding schedule, the animal is placed in the housing-unit and fed. A food-cup containing 2 gm. of food-tablets is placed against the front wall of the unit between the two windows. The rat is placed in the apparatus for 15 min. a day until all tablets are eaten on two successive days. In the next stage of adaptation, the door to the stimulus-table is opened, and *S* is fed from the food-wells. Each well is baited with an equal number of tablets, and the rat is allowed 15 min. in which to consume them. When all tablets are eaten within this period on two successive days, *S* is pretrained with a stimulus-object. During the pretraining period, only one food-well is baited. Initially, *S* is fed one tablet each trial with both food-wells uncovered, the position of the tablet being varied randomly from trial to trial. When the rat takes the reward tablet in 18 of 20 trials within a 15-sec. interval, training with the baited food-well is begun. On successive trials, the object is moved from the back to the front of the food-well in steps of ¼ in. If *S* dislodges the stimulus in 15 sec. or less when the food-well is completely covered, subsequent trials are run with the well covered. Failure to dislodge the stimulus from the completely covered well is followed by another series of trials in which successive fractions of the well are covered. After the rat dislodges the stimulus from the covered well within 15 sec. on 18 of 20 trials, discriminative training is begun.

Sample data. Fig. 2 shows the performance of a group of six adult male rats on a two-choice simultaneous discrimination involving multidimensional objects (junk). These *Ss* were given 40 trials a day by a non-correction procedure until they reached a criterion of 90% correct choices in the first or second block of 20 trials on a training day. On the following day, reversal-training was begun. Fig. 2 presents the data expresssed as percentage correct responses in blocks of 20 trials for original learning and one reversal.

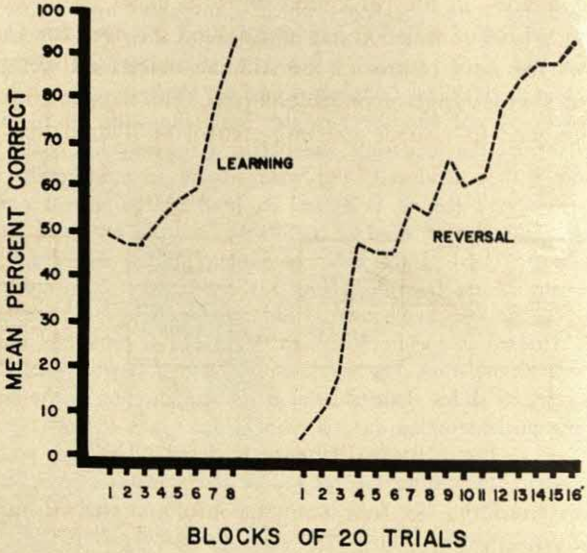


FIG. 2. PERFORMANCE OF SIX RATS ON A TWO-CHOICE SIMULTANEOUS OBJECT-DISCRIMINATION AND ONE REVERSAL

A FLEXIBLE FOUR-FIELD TACHISTOSCOPE

By WILLIAM K. BOARDMAN, Baylor University College of Medicine, and R. C. GUNTER, JR., R. C. Gunter Associates, Charlton, Massachusetts

Increased interest in the perception of form under conditions of multiple or cyclic visual stimulation has highlighted the need for an apparatus which allows for rigid control of the stimulus-objects and yet permits selection from a wide range of possible objects. The tachistoscope described below is designed to provide exacting control of illumination and time-intervals of from one to four stimulus-objects presented in single sequences or cyclically.

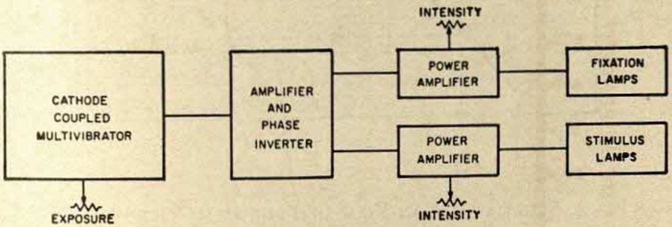


FIG. 1. BLOCK-DIAGRAM OF A TIMER-UNIT

The appartus consists of a viewing box, a series of Gunter Type 200 timers,[1] and a master-control. The timers employ a cathode-coupled multivibrator of a type used in some radar-circuits which allows the bias on the normally off tube to be varied with a 10-turn potentiometer. A similar circuit has been described by Koletsky and Kolers.[2] By switching in four different capacitors, it is possible to obtain a continuously variable duration from below 1.0 m.sec. to 10.0 sec. with an error of no more than 3% for any setting. The output-stage uses power vacuum-pentodes which allow maximal control of the intensity of illumination. The pulse emitted by the activated sources of light is essentially square for all settings in the above range, although the timer-circuit can be modified to produce a pulse of any desired form. A block-diagram of a timer-unit is shown in Fig. 1.

A plan-view of the viewing box is shown in Fig. 2. Each field is 43 in. from

---

* The research reported in this paper was supported in part by grants from the U.S.P.H.S. (M-1121 and M-07374) and the Hogg Foundation.

[1] These timers were originally designed for individual use with a Gunter Type 103 two-field viewing box and are interchangeable for this purpose.

[2] H. S. Koletsky and P. A. Kolers, A multi-field electronic tachistoscope, this JOURNAL, 72, 1959, 456-459.

the viewing aperture and is illuminated by two standard 4-w. fluorescent tubes. The illumination of each field is controlled by the associated timer. The half-silvered mirrors are so designed that intensity and spectral transmission are equivalent for each field, consequently *O* is not provided with any cues through sudden change in color. Stimulus-holders are designed to accept standard 8½ × 11-in. plates, but adaptors allow almost any modification, including the use of a movable strip of paper or film or if three-dimensional objects. The interior of the viewing box is a uniform dull black, and baffles are used to eliminate virtually all stray light.

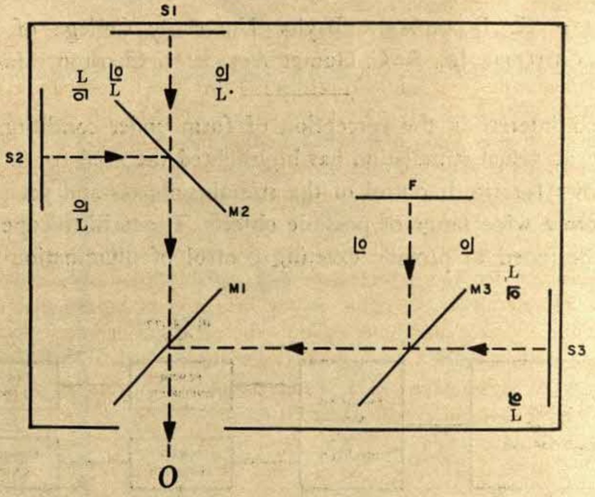


FIG. 2. SCHEMATIC PLAN-VIEW OF FOUR-FIELD VIEWING BOX
L = 4-w. fluorescent lamps; M1, M2, and M3 = Special 'half-silvered' beamsplitter mirrors; S1, S2, and S3 = stimulus-fields; F = fixation-field; *O* = observer; internal baffling, not shown, prevents light from scattering inside the apparatus.

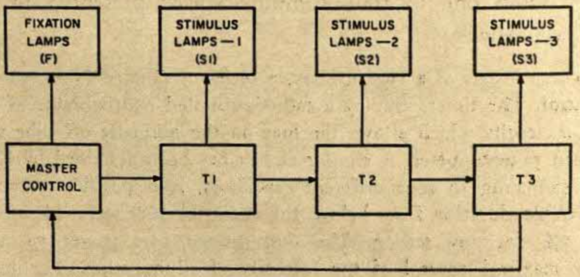


FIG. 3. BLOCK-DIAGRAM OF THE OPERATION OF THE FOUR-FIELD SYSTEM
T1, T2, and T3 are standard Gunter timers, Type 200

The operation of the tachistoscope is diagrammed in Fig. 3. The timing units are interchangeable, and an input-output modification allows them to be used in sequential arrangement with the master unit for multi-field presentation. The master-unit regulates the intensity and duration of the field chosen as the fixation-field. Its circuit includes a pair of thyratrons in a back-to-back arrangement. The thyratrons are used as a switch in much the same way as a bistable multivibrator, but are capable of carrying much heavier currents.

Actuating a noiseless switch on the master-control turns off the fixation-field (F) and initiates a complete sequence of exposures, by triggering timer No. 1 (T1) which illuminates the first stimulus-field (S1). At the end of the exposure controlled by T1, it turns off S1 and sends a trigger pulse to T2, etc. T3 triggers the master-control and F is again illuminated, ending the sequence. Cyclic operation can be substituted for the above manual activation by by-passing the master-control and F such that T3 triggers T1, or by substituting an additional timer for the master-control to obtain a four-field cycle.

The inter-stimulus interval (ISI) of the trigger pulse is 4 µsec., a completely negligible duration for visual presentation. A variable ISI may be obtained in two ways. First, the circuit to the lights of any field may be broken that the field remains dark and its associated timer controls the period of darkness. Secondly, additional timing units may be introduced between those which are used to illuminate the fields of the viewing box. This latter technique provides an almost limitless number of combinations of exposure and ISI.

APPARATUS NOTES

A BINARY PERCEPTUAL MAZE

A common difficulty in developing tests of intellectual functions is that of deriving a series of items for which criteria of difficulty can be varied systematically. The present paper describes the development of a perceptual maze-test in which the individual patterns are free from subjective bias on the part of the designer and in which the main sources of difficulty can be varied independently of each other. As the number of variations in pattern which can be created by use of a simple basic format is almost infinite, movable type which greatly facilitates the construction of new patterns has been developed and is also described.

The test in its present form has evolved from patterns similar to those illustrated in Figs. 1A and 1B.[1] Each pattern has a lattice background of paths or tracks. At a variable number of intersections, dots are superimposed. *S* is required to find in each pattern a pathway from bottom to top, or top to bottom, which passes through the largest number of dots. There are two restrictions. First, *S* must keep to the paths or tracks, and not cut across from one path to another. Secondly, at any junction the path chosen by the *S* must continue forward, *i.e.* it may fork left or right, but must not double back. The criterion of success employed in studies so far undertaken has been the achievement of a correct solution within a certain specified time.

The test-material in its original form was used by one of us in a study of patients undergoing neurological operations. It was found to be sensitive to injury of the frontal and temporal lobe. Scores on the test correlated positively with post-operative social adjustment, scores on tests of verbal and performance intelligence, and negatively with chronological age.[2] Benton has recently confirmed the sensitivity of the test to brain-damage.[3] A group-version of the test has been given to normal *Ss* and found to have a high reliability (test-retest correlation with identical version, $r = 0.89$, test-retest correlation with an equivalent version, $r = 0.81$). Correlation with tests of verbal and non-verbal intelligence ranged between 0.37 and 0.46.[4]

These preliminary results supported the view that this test reliably samples some aspect of intelligent behavior not adequately sampled by conventional tests of intelligence. It has been suggested that this aspect might be the size and complexity of the perceptual units which individuals can effectively handle.[5] Further development of the test seemed justified if the perceptual complexity of the material could be simply and systematically varied. The design of a fully binary or 'V' version satisfied most of the theoretical requirements and the development of movable type has provided a solution to the technical problem. Two basic types of units—crossing lines with and without dots—have been designed, either of which can be easily modified to form edge-pieces and with these any required pattern can be assembled. In designing this type, greater perceptual uniformity was achieved by making the apical angles 60°. Each choice-point is then equidistant both horizontally and diagonally from the next.

---

* Our thanks are due Dr. E. A. Carmichael, Mr. Caslon Chitson, Miss Margaret Eeuwens and Mr. D. N. Lee for their assistance in the development of this test.

[1] Alick Elithorn, A preliminary report on a perceptual maze test sensitive to brain damage, *J. neurol. neurosurg. Psychiat.*, 18, 1955, 287-292.

[2] Elithorn, *op. cit.*, 289.

[3] A. L. Benton, Personal communication, 1960. Professor Benton's study was part of an extensive investigation of the differential impairment of psychological functions by focal brain-lesions, the details of which will be published later.

[4] Alick Elithorn, Myfanwy Kerr and Joy Mott, A group version of a perceptual maze test, *Brit. J. Psychol.*, 51, 1960, 19-26.

[5] Elithorn, *op. cit.*, 291.

A 'V' version made from this type and illustrated in Figs. 1C and 1D has several

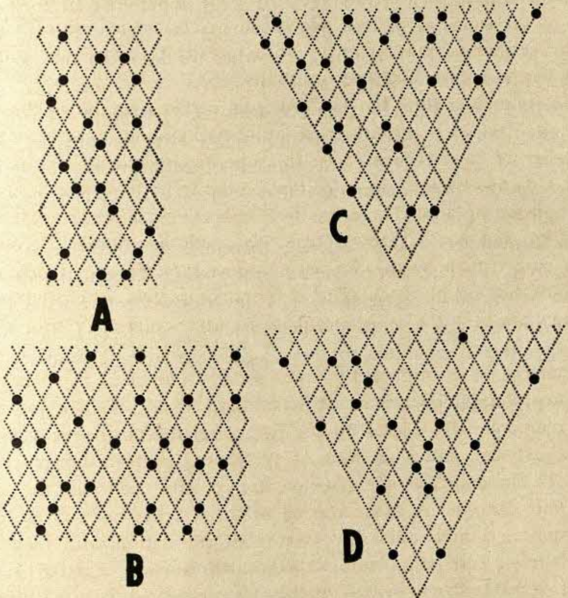


FIG. 1. EXAMPLES OF PATTERNS USED AS PERCEPTUAL MAZES

theoretical advantages over the rectangular version. The fully binary structure itself imposes a consistency on the task which was not present with the rectangular format. In the latter, the sides of the pattern offered target-points at which there was no choice, a reëntrant path being obligatory. In the 'V' version, the previously uncontrolled variable—shape—is held constant, all patterns having the same triangular form. While reducing the number of uncontrolled variables, this restriction has the disadvantage that the number of possible pathways cannot be varied independently of their length. The three variables of size, saturation, and pattern (position of dots in relation to background) can, however, be controlled and vareid systematically.

In this version, size of pattern can be expressed in many ways. First, as the length of pathway (number of steps in the lattice); secondly, as the total number of possible pathways; third, as the number of target-points, comprising all points of intersection except the first which, being common to all pathways, cannot be either selected or rejected; finally, as the number of choice-points, defined as all points of intersection excluding the final row, which provides no choice.

If the rows of the lattice are numbered 0, 1, 2, 3 . . . . n, starting at the vertex, the above can be expressed: (1) Length of pathway $= n$; (2) Total number of pathways $= 2^n$; (3) Number of target points $= N = [((n+1)(n+2))/2]-1$; (4) Number of choice points $= [n(n+1)]/2$.

The saturation of any given pattern is defined as the proportion of target-points at which dots are sited: N_t/N, where $N_t =$ number of dots. Where the arrangement of dots is random, saturation corresponds to the probability (p) of a dot occurring at any particular target-point.

The third variable, which is obviously of great importance, is the pattern of relationships among the dots. In the earlier rectangular versions, patterns were built up according to the designers' ideas as to what would make the item difficult. A correct pathway involving several changes of direction, for example, might be expected to provide more difficulty than one whose dots were in a straight line. 'Red herrings' in the form of a concentration of dots away from the correct solution might be expected to distract *S*'s attention. The possibility of developing patterns by a purely random technique means that sources of difficulty may be revealed which have not been foreseen.

The main advantages of this material, therefore appear to be as follows: (1) The complexity of the task can be varied over a wide range and several sources of difficulty can be varied independently. (2) Individual items can be generated statistically and bias on the part of the designer eliminated. (3) The test-items, answers, and hypotheses about the sources of difficulty are all suitable for treatment by computer.

The Royal Free Hospital
London, England

ALICK ELITHORN
MYFANWY KERR
DAVID JONES

NOTES AND DISCUSSIONS

EXPERIMENTAL AND PHYSIOLOGICAL PSYCHOLOGY: HISTORICAL RELATIONS

In a previous note,[1] we traced the rise of experimental work in psychology, as catalogued by Rand, from 1797 to 1874.[2] We suggested that the impression of the experimental development in psychology might be modified if the items classified by Rand as Experimental and Physiological were reclassified. Such a reclassification was made and the results are reported here. The items in that category were reclassified as (*E*), mainly experimental; as (*P*), mainly physiological; and as (*EP*), belonging to both. To trace the impact of the 'New Psychology,' this study rounds out the century by including the years 1875-1896.

The attempt to reclassify the studies in Rand's "Experimental and Physiological" category shows at once why he was content with one overall list in this field. This category includes items which are general, or might be either *E* or *P* from their titles, or are of interest at one remove from the main subject matter. Nor is all physiological psychology the same: the meaning of this term differs widely among the authors. A final sub-classification could be arrived at only by reading the entire content of the entries and rating them on suitable scales from *E* to *P*. For a working classification, however, titles and general background knowledge are enough, bearing in mind that the *EP* or compromise sub-class must remain large.

From 1797 to 1874 inclusive, Rand lists 140 entries. The sub-classification is: *E* 29 (20.7%); *P* 89 (63.6%); *EP* 22 (15.7%). From 1875 to 1896, there are 371 entries. Sub-classified, these are *E* 152 (41.0%); *P* 132 (35.6%); and *EP* 87 (23.4%). Taking over-all figures, the whole century, 1797-1896, covers 511 entries, of which 181 (35.4%) are *E;* 221 (43.2%) are *P;* and 109 (21.3%) are *EP*. The numbers and percentages of entries per decade are now tabulated.

Thus, if one's interest is in the over-all development of experimental-physiological work, before and after 1874, the year marked by Wundt's

---

[1] J. A. Cardno, The experimental upsurge: 1797-1874, this JOURNAL, 75, 1962, 499-501.

[2] Benjamin Rand, Bibliography of philosophy, psychology, and cognate subjects, in J. M. Baldwin (ed.), *Dictionary of Philosophy and Psychology*, Vol. 3, Part 2, 1905, 950-964.

Grundzüge and Brentano's *Psychologie vom Empirischen Standpunkt,* he will find that the output for the 22 yr. after that date is 265% (371:140) of that for the 78 yr. before it. The composition is different, too. *E* has almost doubled in proportion to total output; *P* has almost halved. The rise in *EP* may reflect uncertainty in part, for it is peripheral to the writer's special period of interest. It reflects also some increase in critical discussions applicable to both *E* and *P*, and in general texts. The impression is, however, that uncertainty leads to understatement of the experimental upsurge, if anything.

The century as a whole shows rises in production in this field from decade to decade. This increase more than doubles from 1857-1866 to 1867-1876, and again from 1877-1886 to 1887-1896, thus supporting the view that experimental psychology in the broad sense really got going from about 1860 onwards. The proportional increase in *E* is much more marked than in *P*. Remove the last two decades in *E*, and only 18.3% of the century's production is left. In *P*, these same two decades cover only 54.8% of production.

TABLE I

Number and Percentage of Entries per Decade

		Decades: 1797 to 1896									
		97–06	07–16	17–26	27–36	37–46	47–56	57–66	67–76	77–86	87–96
Rand per decade:	No.	2	3	4	9	18	24	31	70	113	237
(511 = 100%)	%	0.4	0.6	0.8	1.8	3.5	4.7	6.1	13.7	22.1	46.4
E per decade:	No.	1	2	1	3	4	3	6	13	36	112
(181 100%)	%	0.6	1.1	0.6	1.7	2.1	1.7	3.3	7.2	19.9	61.9
P per decade	No.	1	1	2	6	13	17	22	38	45	76
(221 = 100%)	%	0.5	0.5	0.9	2.7	5.9	7.7	9.9	17.2	20.4	34.4
EP per decade	No.	0	0	0	0	1	4	3	19	32	49
(109 = 100%)	%	0	0	0.9	0	0.9	3.7	2.8	17.4	29.3	45.0

While figures for particular years are liable to be uncertain, a few qualitative comments are worth making. *E* production is very sparse until 1860, which has two examples: before this there are single works in each of the years 1802, 1814, 1822, 1827, 1831, 1832, 1842, 1846, 1850, 1852, 1855, 1859. The first dual entry occurred in 1839. The striking experimental (*E*) upsurge, however, comes very late. Double figures in *E* are first reached in 1888, then there is a lull from 1889 to 1892, followed by a rise to 23 (1893), 19 (1894), 19 (1895), and 13 (1896). *P* production by years is a less spectacular story. Production is sparse up to 1841, with one entry each in 1806, 1816, 1829, 1831, 1834, 1835, with two entries in 1819 and 1833. From 1841 onwards, entries appear for most years, rising occasionally to four in a year (1843, 1852, 1864). In 1874, there are four entries, in 1876, nine. Thereafter, production is never less than three in a year, but double figures are reached only in 1891, and the highest figures (11 in 1894, 14 in 1895) remain considerably lower than those for *E*.

In sum, then, the impression of great experimental (*E*) development is strengthened rather than shaken. One's inclination to expect experimental developments mainly from 1860 on, derived from variations in writers' uses of the adjective 'experimental,' is confirmed. One regards the physiological (*P*) development as more gradual than the experimental; more like the areas Abnormal and Criminal, and Insanity and Idiocy, which, on Rand's showing, represent standing problems of psychology rather than striking new movements. The growth of *EP*, even allowing for uncertainties, suggests that the increasing volume of discussion *about* experimental and physiological work deserves attention.

In a wider historical context, the increase in *P* production from 1841 is a signpost to the influences at work on outstanding early physiological writers, such as Bain. The late upsurge of *E* raises the question whether the productive influence in Leipzig was the laboratory (1879), the *Philosophische Studien* (1881-1903), or the arrival of American students.[3]

University of Tasmania J. A. CARDNO
Hobart, Tasmania

EFFECT OF THE MODULUS ON ESTIMATES OF MAGNITUDE OF LINEAR EXTENT

Stevens has discussed some of the factors which may affect an observer's ability to make numerical estimates that accurately reflect psychological values.[1] One factor suggested was the modulus—the numeral assigned to the standard stimulus. The present study was undertaken to discover the effect of the modulus upon the estimation of visual length.

Seventeen lines of different lengths, varying from 0.25-7.25 in., were drawn diagonally from one corner of separate 4 × 6-in. index cards. Fifty-seven women, paid to serve as *Ss*, each made two estimates of the length of every line; one estimate in terms of each of two moduli. Nineteen *Ss* made judgments with the standard called '2' and '91,' 19, with '4' and '182;' and 19, with '6' and '273.' The standard stimulus was always the same length, 1.25 in. The order of presentation of the stimuli was randomized within each of the two series.

---

[3] E. G. Boring, *A History of Experimental Psychology,* 1950, 324, 533.

[1] S. S. Stevens, The direct estimation of sensory magnitudes—loudness, this JOURNAL, 69, 1956, 1-25; The psychophysics of sensory function, *Amer. Scient.,* 48, 1960, 226-253.

The logarithms of the estimates of magnitude were plotted against the logarithms of the stimulus-values and straight lines were fitted by the method of least squares. The exponents of the psychophysical functions were given by the slopes of the lines, and were plotted against the logarithms of the moduli in Fig. 1. They vary, as the figure shows, between 1.03 and 0.89. The tendency for larger moduli to give lower exponents was significant at the 1% level. All exponents were lower than the value, 1.1, given by Stevens for length of line. These results may be due to differences in experimental technique, or to difference in the *S*s.

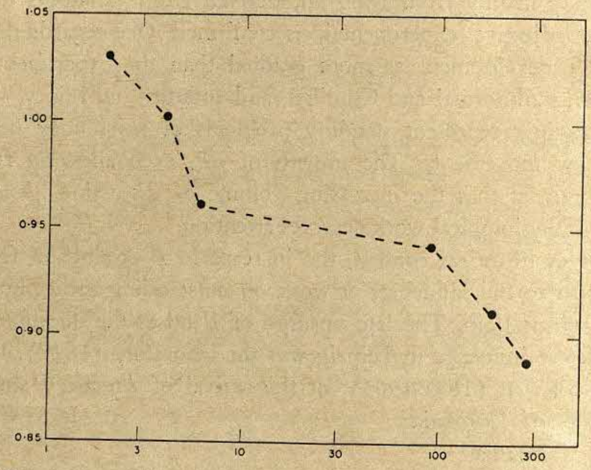


FIG. 1. EXPONENT FOR LENGTH OF LINES AS A FUNCTION OF THE MODULUS

Defence Research Medical Laboratories
Toronto, Canada

RODERICK WONG

THE FIFTY-FIFTH ANNUAL MEETING OF THE SOUTHERN SOCIETY FOR PHILOSOPHY AND PSYCHOLOGY

The Southern Society for Philosophy and Psychology held its fifty-fifth annual meeting at the Hotel Seville in Miami Beach, Florida, on April 11-13, 1963. Approximately 175 members, friends, and guests attended. Hosts were the Departments of Philosophy and Psychology at the University of Miami.

The program consisted of 5 symposia or special reports in psychology and 2 in philosophy, 24 papers in each of the two fields, and 3 joint sessions.[1] Two Junior Awards were given: one for a paper presented by Richard A. Smyth (University of North Carolina) on "Spinoza's fourteenth proposition," and the other for a paper presented by Leonard E. Jarrard (Washington and Lee University) on "Effects of hippocampal ablation on bar-pressing for food." These awards are citations for general excellence; they are given annually for the papers of highest merit presented by members who have not yet received the doctoral degree or who have received it within the preceding five years.

At the first of the three joint sessions, John Paul Nafe (Florida State University) introduced Max F. Meyer (Springfield, Va.) who presented an invited address on "The present status of the theory of hearing." William M. Hinton (Washington and Lee University) served as Chairman of the subsequent two joint sessions; as Past President of the Society, in the first of these he introduced Henry King Stanford (President, University of Miami), who presented a "Message of welcome," and Leroy E. Loemker (Emory University), who delivered the presidential address entitled, "The case of the emperor's clothes." The third joint session was an interdisciplinary symposium on sensation; the invited principal speaker was Charles Hartshorne (University of Texas), and the respondents were Karl E. Zener and Charles A. Baylis (both of Duke University).

Eighteen philosophers and thirty-two psychologists were elected to full or associate mmebership during the annual business meeting. Four former presidents of the Society were elected Honorary Members for Life—a class of membership which represents the highest honor the Society can bestow, and election to which attests to the high regard with which the awardee is held in the minds and hearts of his colleagues. Those elected, and year of their tenure as president, were: A. S. Edwards (1927), University of Georgia; Marten ten Hoor (1938), University of Alabama; Fritz Marti (1942), Berglihof School, Ohio; and Karl M. Dallenbach (1954), University of Texas.

The following officers were elected to serve during 1963-1964: Dan R. Kenshalo (Florida State University), President; Charles Hartshorne (University of Texas), President-Elect; and Leroy E. Loemker (Emory University), Past President. Earl A. Alluisi (Lockheed-Georgia Company) and Leland E. Thune (Vanderbilt University) continue to serve as Secretary and Treasurer, respectively. E. M. Adams (University of North Carolina) and Malcolm D. Arnoult (Texas Christian University) were elected to the Council as representatives of philosophy and psychology, respectively. Other members of the Council are Paul S. Siegel (University of Alabama), Edward A. Bilodeau (Tulane University), John J. Compton (Vanderbilt University), and Herbert C. Sanborn (Brentwood, Tenn.).

---

[1] A copy of the program, including a syllabus of Dr. Meyer's invited address, has been placed on file with the American Documentation Institute (ADI) as Document No. 7566. Photocopies or 35-mm. microfilm copies may be obtained by remitting $1.25 to ADI, Photoduplication Service, Auxiliary Publications Project, Library of Congress, Washington 25, D.C.

The members accepted the invitation of the University of Kentucky to hold the Society's fifty-sixth annual meeting at Lexington, Kentucky, on March 26-28, 1964.

Lockheed-Georgia Company
Marietta, Ga.
EARL A. ALLUISI

THIRTY-FOURTH ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The thirty-fourth annual meeting of the Eastern Psychological Association was held April 11-13, 1963 at the Statler Hilton Hotel in New York City. Total registration at the meeting was 2,745 persons. Of this number 1,344 were members of EPA, 216 were new members who joined the association at the meeting, and 1,185 were guests. The present membership of EPA totals 3,666.

Raymond D. Hedberg was chairman of local arrangements, assisted by Robert Tanofsky, Richard S. Barrett, Waldan D. Setzfand, Samuel H. Cleff, and Saul Gellerman. The Program Committee consisted of John L. Kennedy, Chairman, Henry P. David, and Leonard Ferguson. The committee scheduled 325 papers (presented in 54 sessions), 10 symposia, 6 special sessions, and 4 special group meetings.

The majority of individual research papers related to topics in experimental, comparative, and physiological psychology. Titles of paper-reading sessions in these areas were as follows: operant behavior, animal learning, reinforcement, avoidance behavior, conditioning, discrimination learning, perceptual learning, sensori-motor coordination, studies in reading, paired-associate learning, human learning, cognitive processes, verbal behavior, decision processes, perception and engineering psychology, visual perception, vision, sensory processes, sensory measurement, sensory deprivation and stress, human motivation, animal motivation, animal behavior, psychopharmacology, physiological and neuropsychology.

Three program sessions were devoted to developmental and child behavior. Social psychology was covered by sessions entitled social attitudes, social behavior, social perception, and group behavior. Papers in clinical research were grouped under clinical evaluation, schizophrenia, therapeutic processes, clinical judgment, and projective techniques. Other sessions related to personality and aesthetics, industrial and educational psychology.

Brief titles of symposia were "Symptom treatment in psychotherapy," "Crew performance validation in manned space flight," "Cognitive functioning in elderly people," "Relationship of eye movements to dreaming and waking imagery," "Choice behavior in animals and humans," "Selected aspects of contemporary psychology in the U.S.S.R.," "Pain and human behavior," "Industrial training," "Representation of psychological theory by means of list-processing languages," and "The Peruvian hacienda peon."

Louis M. Herman, top winner of the First Annual Creative Talent Award, was an invited speaker at the meeting of Psi Chi. An invited address with the title "Perceptual recognition as problem solving" was given by Jerome S. Bruner. Richard L. Solomon delivered the annual Presidential Address entitled "Punishment."

During the business meeting, Charles N. Cofer was announced as the President-elect and Jerome S. Bruner and Joseph V. Brady as new members of the Board of Directors. The membership unanimously endorsed a proposal to protest bills before the current session of the U.S. Congress which would unduly restrict scientific research with animals, but to support legislation on animal care which would call for self-discipline by committees of scientists. In the budget for 1963-64, an expenditure of $1,500 was approved to enable the Program Committee to invite distinguished scientists to participate in the forthcoming Annual Meeting.

The 1964 meeting of the association will be held at the Bellevue-Stratford and Sylvania hotels in Philadelphia, April 16-18.

Flushing, New York MARVIN A. IVERSON

THIRTY-THIRD ANNUAL MEETING OF THE ROCKY MOUNTAIN PSYCHOLOGICAL ASSOCIATION

The thirty-third annual meeting of the Rocky Mountain Psychological Association was held in Cody, Wyoming, May 9-11, 1963. Approximately 140 persons attended.

The program included both papers and symposia. The invited address by Dr. Harold McCracken, Director of the Whitney Gallery of Western Art, entitled "Western art as a part of our historical heritage" was one of the highlights of the convention. The presidential address by Dr. Earl Swartzlander was called "Son o' riches."

Officers for the coming year are: President, Merrell Thompson; President-elect, John Cambareri; Secretary, John Thompson; and Treasurer, Richard Lake.

The 1964 meeting will be held in Salt Lake City, Utah, with William Brown serving as convention manager.

University of Denver WILBUR C. MILLER

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Emotion and Personality: Vol. I, Psychological Aspects; Vol. II, Neurological and Physiological Aspects. By MAGDA B. ARNOLD. New York, Columbia University Press, 1960. Pp. Vol. I, xiv, 296; Vol. II, xvii, 430. $7.50 each volume.

In these volumes Magda B. Arnold presents a comprehensive analysis of emotion. The work is scholarly, well documented by extensive bibliographies, clearly written. It is an important contribution to the literature of emotion which should be taken seriously by psychologists, psychiatrists, physiologists, and others concerned with the nature and role of emotion in human life.

Dr. Arnold aims to present a theory of emotion that is both balanced and complete. A *balanced* theory must explain the nature of emotional experience, how it differs from unemotional experience, how emotions are aroused, what bodily changes occur in emotion, and how emotions are related to specific goal-directed actions. A *complete* theory must further show—psychologically and physiologically—how one emotion differs from another, how feelings and emotions are related, what neurophysiological mechanisms mediate the experience and expression of emotions, and the significance of emotion for personality integration. To produce a theory of emotion that is both balanced and complete is a big order!

In a historical survey of scientific theories of emotion Arnold dates the beginning of the modern era with James rather than Darwin on the ground that Darwin's basic principle of *serviceable associated habits* implies the inheritance of acquired characters. Arnold reviews psycholanalytical theories; the psychological theories of George Dumas; the behavioral theories of Watson, Hull, Skinner, Tolman; the attitude theory of Nina Bull; the energy release theory expressed by Meyer and developed by Duffy; the motivational theory of Leeper; the eclecticism of Young; the existential reaction of Sartre; and the functional theories of Gemelli and Michotte. She feels that the theories of Gemelli and Michotte are balanced though not complete.

Her own theory can best be described as a special emphasis within a comprehensive view. The emphasis is upon *appraisal* of the object or situation by an individual. Appraisals are relative to the weal or woe of an individual—to the way in which a situation affects *me* and *my welfare.* In the past, psychologists have studied the sequence perception→expression→emotion rather than perception→appraisal→emotion which alone can explain the conditions necessary for emotional arousal. Animals appraise directly and intuitively what they see, hear, touch. Humans appraise not only by sensing but by remembering, reflecting, reasoning, understanding.

Arnold describes a separate neural system for appraisal. This estimative system receives relays from the sensory system and includes afferent connections to the brain stem reticular formation, intralaminar and midline thalamic nuclei, and cortex of the limbic lobe. Following Pribram, MacLean, and others, Arnold considers in detail the neural basis for evaluations of equilibrium, sounds, sights, tastes, somesthetic impressions, smells, action and bodily movements, eating, and speaking.

In discussing the systematic status of feeling, Arnold reviews introspective studies beginning with Wundt and Titchener. She concludes that feeling is an experience *sui generis,* not to be identified with organic sensation nor with any other sensory process; nor is feeling a construct. Pleasantness and unpleasantness are direct unlearned experiences. She regards P and U as immediate, primitive, biological appraisals that an object or situation is good or bad. With man these immediate affective evaluations are supplemented by memory and reflection.

To the reviewer it seems that appraisal is fundamentally a *cognitive* process and that Arnold's view leans towards the judgmental theory of affectivity. A question can be raised as to whether this view of appraisal is consistent with the view that P and U are feelings *sui generis.* But, in any event, a cognitive appraisal of the situation does make a tremendous difference in human behavior. We recall that James long ago pointed out that meeting a bear in the woods is different from seeing a bear in the zoo!

In classifying basic emotions Arnold recognizes three main dichotomies. The object or situation is appraised as: (1) good or bad, (2) present or absent, (3) difficult or easy to attain (or avoid). These three, indeed, are important distinctions but they yield a logical rather than a psychological or physiological classification of emotions.

Arnold defines emotion as ". . . *the felt tendency toward anything intuitively appraised as good (beneficial), or away from anything intuitively appraised as bad (harmful). This attraction or aversion is accompanied by a pattern of physiological changes organized toward approach or withdrawal. The patterns differ for different emotions.*" (Vol. I, p. 182)

An important part of the work is a group of chapters in the second volume dealing with neurological and physiological processes. Arnold accepts the works of Papez and MacLean as the best basis for describing the brain mechanisms of emotion. She rejects Cannon's hypothesis that emotion is generated in the thalamus or hypothalamus and that it is experienced and expressed when these structures are released from cortical inhibition. Opposed to Cannon's "release" theory is the excitatory theory of Papez. The great contribution of Papez lies in his postulate of a definite area of the brain which mediates the *experience* of emotion and in his sketching of possible neural circuits connected with this area. Papez's theory makes it possible to explain emotion as a *centrally initiated* process. He also held that emotional expressions can be aroused by sensory impulses related to the hypothalamus and cingulate gyrus on their way to the sensory cortex. An important addition to Papez's theory is MacLean's conception of a "visceral brain" with centers in the limbic lobe.

Arnold insists that an adequate account of conscious emotion is an integral part of a complete theory. Behaviorists have ignored consciousness and as a consequence we have the strange phenomenon of physiologists and neurologists propounding theories of emotion (Cannon, Papez, MacLean) while psychologists keep scornfully aloof. Or neurologists and anatomists wrestling with the problem of consciousness while psychologists consider this to be a semantic pseudoproblem (Lashley, Hebb). Hebb rejects consciousness and operates in terms of behavior and hypothetical neural circuits; but Hebb's system becomes consistent, Arnold states, only on the basis of a dialectic which he repudiates.

Arnold's theory of emotion should be classified as *motivational* because she defines emotion as a felt tendency that leads to integrated action following appraisal of the situation. With a *motivational* theory the determinants of integrated behavior are primary; emotional disturbances are secondary. There is an opposed view, however, which holds that affective disturbance or upset is the earmark of emotion and that a concept of emotion would be unnecessary apart from its application to affective disturbances.

The conflict between these two views of emotion becomes acute when Arnold considers emotion in relation to personality organization. Toward the close of the second volume she gives the disruptive concept full play. But instead of identifying emotion with affective disturbance she consistently argues that emotion is a tendency that urges toward integrated, goal-directed action and that emotion may become disturbing when it is intense, excessive or chronic. According to the disruption view, contrastingly, a disturbed affective state *is* an emotion and the underlying dynamic processes are motivational in nature. The difference between these views doubtless lies in the way one regards complex affective disturbances and their determinants.

In considering the relation between emotion and motivation Arnold gives a clear critique of instinct, drive, need, homeostasis, and other motivational concepts. Pleasure, she writes, is not a motive but a by-product of action. She accepts the common-sense view of motive as a *subjective precondition* or tendency rather than an external force or cause. External determinants are not motives even though they do arouse and influence behavior. Thus, Arnold holds a narrow view of motivation and a broad view of emotion rather than *vice versa.*

In a final section she takes a philosophical (even moral and religious) view of the self-ideal, disregarding the empirical findings reported in earlier chapters. The reader notices a radical shift from a physiological frame of reference to something else. The analysis of shame, remorse, repentance, guilt, boredom, human suffering, contrition, self-blame, self-condemnation, and other feelings, reminds one of the writings of Shand and McDougall rather than the careful physiological analysis of previous chapters. And in a final chapter positive human "emotions" are discussed, among them: union with others, infectious emotions, empathy, sympathy, love, admiration, love of beauty, love of God, happiness. In this discussion the concept of emotion becomes practically synonymous with affectivity in general.

There is an appendix that takes up problems of research and measurement. Arnold suggests several constructive approaches. She believes that the best possible test for evaluating human motivation is a story-telling test. In telling a story the individual is forced to fit emotion into action, and to do it in a way that agrees with his attitudes and convictions. Emotion-in-action means emotion controlled and directed, just as imagination-in-action and memory-in-action mean directed and controlled imagination and memory. Goal-directed action, Arnold states, includes both emotion and imagination, in fact, includes every single psychological function. The performances from which intelligence is inferred include little emotion and a great deal of imagination and memory, while sustained purposive activity depends primarily on a man's motives and well-directed "emotions."

These volumes contain important materials. In Vol. I there is an excellent review of the classical introspective studies of feeling. In Vol. II there is a wealth of physiological and neurological detail, well illustrated by diagrams of neural tracts and circuits, and many specific suggestions concerning the bodily basis of particular emotional patterns. And throughout both volumes, as noted above, there is an important emphasis upon appraisal of the situation as a prime condition of emotional arousal.

University of Illinois PAUL THOMAS YOUNG

Visual Problems of Color—Symposium. By NATIONAL PHYSICS LABORATORY (Great Britain). New York, Chemical Publishing Co., Inc. 1961. Vol. I Pp 395; Vol. II Pp 367. $8.50.

"This symposium of two volumes, on an important branch of biophysics, contains a collection of articles by over forty world famous scientists, from eleven countries, who are reporting on their own researches investigating every aspect of visual problems of color." This claim by the publisher is modest enough, but the implied catholicity of the reports may be questioned and the sectarian character of a majority of the papers is evident. "Volume I covers retinal chemistry and the physiology of vision; visual pigmentation . . . brightness matching and color matching." "Volume II covers subjective color measurements, temporal effects and defective color vision; electrophysiological aspects . . . and color theories." As even this brief listing indicates, the researches reported deal almost exclusively either with what may be the biophysical and biochemical properties of the photo-sensitive mechanism in the eye, or with one major aspect of the stimulus for vision, i.e. tri-stimulus mixture. Very little attention is given to visual data or color judgments. One looks in vain for experimental studies of color discrimination in terms of stimulus variables such as wavelength, energy level, area and duration and their effects on qualitative color relationships, color scales and comparable psychological problems. Indeed, the emphasis is so much on stimuli and mechanisms that it seems more appropriate to entitle the symposium "Physical and physiological correlates of color vision."

Within this scope, the symposium presents the authoritative word on many points, as of the date of presentation. However, it must be noted that there has been a lapse of some four years—1957 to 1961—between presentation and publication. In the meantime, some of the pronouncements have become classic and some have been superseded by further work, often by the same author. Even in the latter case, it is valuable to have at hand the earlier stages in the development of an idea.

It is beyond the scope of a review to analyze and evaluate the contribution of every one of the 41 reports. We shall attempt, rather, to point out some of the things that will be of interest to readers with a psychological point of view.

The memorial address on "Retinal Chemistry and the Physiology of Vision," given in honor of Professor Selig Hecht by George Wald of Harvard University, presents an excellent summary of our knowledge about retinal chemistry and visual pigments, including both his own work and an extensive bibliography of 104 references.

For human vision, Wald's comparisons, after various corrections for ocular transmission, between rhodopsin in rods and scotopic sensitivity show excellent agreement. He speculates on the implications of all this work for human cone pigments and makes some predictions as to their eventual characterization based upon their probable chemical composition and upon the data of Stiles and of Rushton.

The symposium also included papers by many authorities working on a physiological explanation of color vision by chemical and electro-physiological means. In this regard, the volumes provide a ready reference source for a number of famous techniques, most of which are too new to be found in single references elswhere. Rushton's measurements of human retinal pigments, by analyzing the light reflected from the fovea, are the subject of one such paper. Both the techniques and the data obtained from color normals, protanopes, and deuteranopes are given. Other techniques for analyzing the response of the retina to light and data on the absorption of visual pigments, are given by Morton and Pitt, Dartnell, Hubbard, and Denton.

A brief paper by Granit introduces the problems and techniques of micro-electrode methods. The work of MacNichol, Macpherson and Svaetichin on the spectral response curves found in fish retinae gives examples of the single luminosity L type and of the opponent color response curves. This section also includes a report on dominator-modulator responses by Donnor and a summary of electro-retinogram data by Brindley.

Further exploration of tri-stimulus mixture functions has been the objective of at least 14 reports. Half of these were contributed by workers from Russia. By for the most important data are those reported by W. S. Stiles of the National Physical Laboratory. The importance of his work stems from Dr. Stiles' status in the field of colorimetry, the superb equipment at his disposal, and the large number of observations made by him and his associates. Unfortunately, the published presentation has not achieved a high degree of clarity. For example, results are presented in terms of cm^{-1} values, which are related to wave frequency, instead of the familiar wave lengths in mμ. Of course, transformation can be made readily if one has a calculator at hand, but since neither physical measure has a simple functional relation to color discrimination, wave frequency has no advantage for the study of color over wavelength. Comprehension of the paper is further impeded by the omission of titles and legends from all of the 13 figures used to present results.

An innovation in the literature of tri-stimulus mixture is the introduction of a new primary at 580 mμ and the shift of the 445 mμ primary to 470 mμ. The purpose of these changes in operational primaries is to improve the spectral matches in the affected regions by reducing the amount of desaturation of the spectral standards. The new primaries are in accord with the established unique colors and the known laws of complementary and non-complementary color mixture. It would be interesting to one concerned with color judgments to explore the results from this angle but such study is precluded by the mathematical transformation of the data into the hypothetical primaries of orthodox tri-stimulus theory.

A detailed comparison of some aspects of defective color vision with normal was presented by Graham and Hsia. The study, was facilitated by the availability of an observer with one eye color defective and the other normal. Luminosity curves of each eye served to classify one as normal and the other as deuteranopic. Color mixture data serve to differentiate the two eyes as respectively trichromatic and dichromatic. Point-by-point comparison of the colors of the spectrum to the two eyes gives an informative picture of color vision with the defective eye. All colors of the spectrum appear to the defective eye as the same color as one or the other of two narrow bands to the normal eye; viz. 465-475 mμ and 560-580 mμ. Other aspects of color defect are reported in four other papers which include descriptions of two new instruments.

One turns to the report by Hurvich and Jameson on "Further development of a qualified opponent-colours theory" with the hope of finding a more psychological treatment of color problems. The result is both rewarding and disappointing. Valuable data are cited that bear on the problem of blue sensitivity at the center of the field of vision. This is a field in which experimental data have long been needed. For example, Hurvich and Jameson show the changes that occur in tri-stimulus equivalents when size of the foveal stimulus is reduced from 2° to 25′. Similarly, shifts in chromaticity are given resulting from several levels of adaptation. Since the use of the CIE chromaticity diagrams presents the findings in mixture equivalents rather than in color judgments, it would have been interesting and informative if the data were plotted, also, in terms of the "opponent-color theory." One could judge, then, the extent to which the latter gives information about color judgments as well as stimulus equivalents. It still is not clear from the résumé of the "opponent-color theory" what mechanism is provided for seeing black, the opponent of white. Luminosity or physical intensity is not the answer.

A most welcome aspect of the symposium is the inclusion of a sizeable Russian contribution. The ten papers which they submitted give us a sampling of their thinking in this field and the work they are doing, information which many of us lack. The ignorance is not mutual, however, for nearly two-thirds of their references are non-Russian.

The brief biographies given for each participant indicate that 10 of the eleven Soviet investigators, are trained in physics; the remaining man's interest is listed as physiological optics. Apparently, then, the investigation of color vision is considered to be more physical than a physiological problem; and as we know. psychology has an entirely different connotation in Russia than in the United States.

It is interesting to see that the problems attacked in the Soviet Union are similar to those being studied in the West. The reports often do not give descriptions of the methods and procedures in as much detail as is the practice in western experimental reports. One looks in vain for reference to any of the classical psychophysical methods, but this is to be expected since the investigators are physicists.

It is to be hoped that more such symposia will be held, that the Russians will continue to attend and that we learn more of their progress.

As stated in the beginning, no attempt has been made to summarize all of the details of experimentation and theorization presented in the Symposium, but we believe we have given a sampling of the material that will tell a reader the technical nature of the topics covered and thus perhaps intrigue him to delve deeper.

U.S. Naval Medical Research Laboratory
New London, Conn.

FORREST L. DIMMICK
JO ANN S. KINNEY
S. M. LURIA

The Psychology of Deafness. Techniques of Appraisal for Rehabilitation. By EDNA SIMON LEVINE. New York, Columbia University Press, 1960. pp. xii, 383. $7.50.

Historically, psychological studies of deafness and deaf persons in this country have been concerned primarily with pragmatic aspects of educational and rehabilitative processes, and their results have been communicated, notably in the pages of the *American Annals of the Deaf,* to an audience consisting largely of educators of the deaf. Edna Simon Levine's book continues in this tradition, as indicated by its sub-title. It departs from tradition, however, in being written for professional psychologists. Thus, in a sense, the book reverses previous roles; where psychological research was formerly brought to the people who knew "deafness," Dr. Levine sets out to bring her knowledge of deafness and deaf people to those who know psychology.

The principal emphasis of the book is consistent with the author's stated purpose, namely, "to provide guides for the psychological appraisal of adults and children with impaired hearing" (p. 3). The "guides" are most specific in Part Three: "Psychological Examination and Evaluation of Persons with Impaired Hearing." This section describes the special techniques of history-taking, interviewing, and testing required by the communication disabilities of deaf subjects, and it provides a comprehensive list of achievement, intelligence, personality, and aptitude tests which can be used with them.

The psychological examiner inexperienced in testing the deaf often finds it difficult to appreciate the stultifying influence of sereve hearing loss upon language development and usage. The pervasive effects of language deprivation on examination procedures and test interpretation are carefully detailed by Dr. Levine. (Dr. Levine omits mention of the observation by others, notably Pintner and Hiskey, that deaf children have a poor sense of the passage of time and that timed tests may therefore underestimate a deaf child's ability.)

Despite these difficulties Dr. Levine makes it clear that with the observance of appropriate precautions clinicians can usually enjoy the comfort of verbal communication with deaf subjects and utilize tests and test procedures with which they are already familiar. The lists of tests in Chapters 8 (adults) and 9 (children) and in Appendix H comprise a library which ought to satisfy any of the routine needs of "appraisal for rehabilitation."

Part Three, which represents the book's sub-title, includes 167 of the 308 pages of main text; the subject-matter of the main title is presented in the 61 pages of Part One: "The Implications of Hearing and Impaired Hearing." The disparity in the quantity of discussion devoted to the latter topic unfortunately entails a comparable disparity in quality.

A major deficiency of this section is its lack of a clear-cut definition of what "deafness" is. On the first page of her first chapter Dr. Levine refers the reader to Appendix B: "Definitions and Classifications of Hearing Impairment," which contains six categories of definitions drawn from at least three different sources and based on widely differing frames of reference. She does not indicate, however, which of the definitions there provided actually guides her own thinking, and the reader apparently must decide for himself whether he should adopt the criteria of otology, of audiology, of pedagogy, or, perchance, of psychology (category 6, p. 314, "According to the Type of Psychic Involvement").

Dr. Levine does eventually define the term "deaf" to mean "onset of severe hearing loss before establishment of speech and language; or other types of hearing loss with education acquired mainly in schools for the deaf" (p. 128). This definition is given some fifty-three pages after Part One has been completed, and even there it is relegated to a footnote.

Rather than resolving previous ambiguities, it compounds them. The emphasis on age of onset is much more specific than the one definition in Appendix B which includes this concept. And the idea that a child is deaf or not according to where he attends school is unique. (Its ironic implication that "deafness" may be produced by a certain type of educational experience is not entirely without foundation.)

Admittedly, deafness is difficult to define with precision. A definition which couples language disability with hearing loss may be most meaningful to educators, but it is an unnecessarily narrow delimitation of what constitutes auditory disability. There is reason to believe that this unidimensional concept will give way to a more global view: that deafness is the total constellation of behavior characteristically associated with loss of hearing. Deafness does not just *have* a psychology, it *is* a psychology. The psychological study of deafness—or better, of hearing impairment—therefore should encompass all ages, all times of onset, all degrees of hearing loss and of speech and language disability, all educational environments.

Dr. Levine implicitly recognizes this by including in Part One a chapter on "Progressive and Sudden Hearing Loss" which deals with the psychological impact of hearing loss in so-called "hard of hearing" persons, whom she categorizes as "(a) the progressively deafened and (b) the suddenly deafened in adulthood" (p. 56). Why she chose to exclude from this discussion (a) the congenitally "hard of hearing" and (b) the suddenly deafened in childhood is not clear. In any case, she apparently feels that the psychology of deafness is not limited to those people whom she defines as deaf.

This chapter, incidentally, provides a striking example of the paucity of research in this area. Of the 39 references cited by the author only 9 are dated later than 1950; 15 of them are dated prior to 1940. It may be questioned whether the early studies are valid for today's population. Recent developments in otological surgery, in hearing aid design and distribution, in expansion of community services for speech and hearing therapy, in the widespread growth of pre-school training programs for hearing impaired children, all combine to ameliorate the effects of hearing impairment, sudden or progressive, much more effectively than could be done prior to World War II. A redefinition of the psychological consequences of "hard of hearingness" is in order.

As with the definition of deafness, so with the psychology of deafness. In Chapter 2 the author presents a generalized discussion of the "psycholinguistic implications," the "psychosocial implications," and the "implications of variable factors" in cases of "early severe deafness;" but when she gets to the sub-heading "Psychological Studies" (p. 50), she simply refers the reader to "Appendix G: Chronological Listing of Selected Published Psychological Studies of the Deaf and Tests Used." This bibliography includes ninety-two items; the basis for selection is not stated. The findings of these studies are not reported, and only two of the publications listed in the bibliography are cited in the chapter itself.

The discussion of psychopathology in the deaf consists of a "check list of psychiatric signs" (p. 131) comprising twelve lines of type, and one-and-a-half pages of textual material (pp. 245-246). (The former is listed in the book's index; the latter is not.)

The differential diagnosis of auditory and language disabilities in children is discussed at somewhat greater length, but the only specific guidance given the clinician in this respect is a list of behavioral signs of deafness (p. 244f) quoted from Gesell and Amatruda's *Developmental Diagnosis,* 1947. Since no attempt is made to contrast these behavioral signs with similar (or the same) signs as exhibited by the aphasic, retarded, and/or emotionally disturbed children from whom the deaf child must be differentiated, the list is virtually meaningless.

The problems of the multiply handicapped deaf person are not discussed at all.

Part Two, "Psychological Practice in a Rehabilitation Setting" "is discussed in the general context of physical disability rather than that of a single disability group" (p. 78). The two chapters of this section have no specific reference or relevance to the psychology of deafness; they make no unique contribution to this volume.

Part Four, "Preparation for Progress" contains a brief discussion of desiderata for research procedures and an extended list of problem areas in which further research is needed. A foreword was contributed by Boyce R. Williams, Consultant, Deaf and the Hard of Hearing, Office of Vocational Rehabilitation, U. S. Department of Health, Education, and Welfare.

To sum up, the book's discussion of the psychology of deafness is essentially a generalized introduction to an incomplete bibliography. As a "short course" in the psychometry of deaf persons, particularly adults, Dr. Levine has written one of the most helpful guides in print.

Indiana University School of Medicine CARL W. FULLER

Brain and Behavior. Volume I: Proceedings of the First Conference. Edited by M. A. B. BRAZIER, Washington, Amer. Inst. Biol. Sci., 1961. Pp. 433. $7.75.

The inaugural conference on Brain and Behavior, sponsored by the Brain Research Institute, University of California, Los Angeles, was held February 19 to 22, 1961 in Los Angeles. Magoun's introductory remarks set the tone for the conference. In them, he echoed John Locke's philosophical empiricism by pointing to the influence of the external environment on the modalities of somatic sensibility, audition and vision that furnish the material upon which the mind employs itself. Appropriately then, this conference was concerned with analyses of each of the major sense modalities. Assembled to participate in the discussions were internationally noted authorities reflecting a variety of disciplinary backgrounds. In the tradition of the Macy volumes, the papers and discussions are published *verbatim* and no attempt is made to develop an integrated story. Were it not for Tueber's masterful reprise and summary, the reader would face an horrendous task of reconciling differing viewpoints and bringing together recurrent themes. Teuber's summation might well be read as an introduction to the ten formal presentations.

Behavior is given second billing in the title of this conference and the overriding impression is that it comes out a poor second in the text. While brain is scrutinized with the micro-methods of electron microscope and ultra-fine electrodes, behavior is peered at through the proverbial wrong end of the telescope; the techniques applied are woefully gross. Should the reader be concerned primarily with the overt behavior of organisms, he will find precious little in this volume to satisfy his interests. What are some of the behavioral techniques and data described at this conference?

1. Touching the human cornea with a nylon thread and other stimuli evoked a variety of reports such as "you tickled my eye"; electrical stimulation of the exposed dorsal columns of the spinal cord in conscious patients evoked reports of unpleasant sensations: "You would have to feel it yourself to be able to describe it, Doctor." On the basis of these and some fascinating anatomic data, Weddell con-

cluded that all the modalities of common sensibility can be evoked from the human cornea provided appropriate stimuli are used. This conclusion was applied to the skin in general and the free nerve ending is given the remarkable burden of mediating the major part of somatic sensibility. This is an heretical but seemingly inescapable view. What of Meissner's corpuscles, Merkel's discs, the Krause end-bulbs and Ruffini end organs? Weddell claims that a cycle of degeneration and repair takes place in cutaneous nerve fibers and these specialized endings are no more than expressions of frustrated nerve regeneration.

2. Implantation of an additional pair of limbs in the newt (to study the temporal and spatial patterns of the walking response of these supernumery limbs) showed that forelimb segments of the spinal cord move both fore- and hindlimbs in the walking patterns of forelimbs, and vice versa. Hence, Szentagothai concluded that the muscles were unable to convert the spinal segments to their own original specificity as Weiss' principle of myotypic specification might predict. Instead, the specific segments forced inappropriate limbs into their own functional pattern.

3. Administration of two types of anesthetics, barbiturates and urethane, showed the EEG correlates of defensive and feeding conditional reflexes to be differentially sensitive to drugs. In spite of the fact that an anesthetic drug completely blocks activation of the cerebral cortex that maintains the waking state, it does not prevent hunger from activating the frontal cortex nor does urethane block the desynchronizing effects of nociceptive stimulation. Accordingly, three activating effects of subcortical centers are postulated by Anokhin: first, activation maintaining wakefulness; second, activation determining the orienting response (attention); and third, activation determining the nociceptive responses of the organism.

4. Pairing of a click with a puff of air to the cat's face produced changes in the EEG response to the click. Galambos' use of the words "conditioning" and "reinforcement" in relation to this experiment evoked a lively discussion of semantics sometimes verging on the ludicrous. Should one refer to the puff of air as a reinforcement? In clarification, Neff proposed an alternative procedure: "What if you presented shock to an animal's hindleg, followed by a flash of light? Would you say that light is the reinforcing stimulus?" Despite the participation of noted psychologists in the discussion, no one succeeded in clarifying the discussion by reviewing the various meanings of reinforcement including reward, contingency and need reduction. Nevertheless a wise word of caution was expressed first by Teuber, then by Neff: changes in neural activity following the pairings of stimuli in the Pavlovian manner are sometimes described as neural correlates of learning. This is not justified until behavior is examined and is found to bear a relation to these changes in neural activity.

5. Total and subtotal ablation of the auditory cortex and transection of the brachium of the inferior colliculus in cats produced the various discrimination deficits of tonal frequency, intensity, duration, patterns and localization (Neff).

6. There appears to be an electrographic correlation with discrimination of flicker from fusion in cats (Lindsley).

This brief survey of topics does not adequately illustrate the overwhelming preoccupation with presumed electrographic correlates of behavior by the participants of this conference. The thread of discussion involving evoked potentials could be followed through most of the papers: mapping of the thalamus (Mountcastle), detailed analysis of the superficial cortical response in neonatal cats (Purpura), phylogenetic comparisons in electrography (Smirnov) and neuronal physiology of the visual system (Creutzfeldt).

This reviewer found the following electrophysiological issues noteworthy for psychologists:

Though the "specific-non-specific" dichotomy is retained, the boundary between these systems has become less sharp and their interaction a subject of deep discussion. In calling attention to some of his tracings, Galambos remarked: "Notice the extraordinary range of brain locations in which . . . a click . . . spread its effects. . . . Notice also, how closely the responses resemble one another in form. The recording pen writes much the same series of deflections . . . regardless of where the electrodes lay in the brain. . . . One can hardly tell the differences between the activity evoked from habenula, hypothalmus and auditory cortex. . . . Electrophysiologists . . . have encountered difficulties in distinguishing the cortical EEG of a man from that of a rat, and now . . . we must recognize the same frustration when we set out to distinguish spontaneous records from cortical and subcortical areas." How reminiscent of the cytoarchitectonic problem encountered by Bailey and von Bonin and the protests made by Lashley and Clark!

In recent years, the centrifugal control of afferents has become a subject of particular interest to psychologists because an attentional mechanism is implied. What a disappointment then to find that habituation and dishabituation may be a matter of conditioned muscular contractions: middle ear muscles in response to clicks (Guzman-Flores) and pupillary reaction to light (Fernandez-Guardiola). This issue, as of this conference, was still unsettled and Teuber sounded an optimistic note.

The issue of the role of d.c. field effects in the cortex is still quite alive. Morrell reported an experiment in which he applied a small anodal polarizing current to the visual cortex. This made it extraordinarily easy for him to drive a single cortical unit at 3/sec. flash frequency. If such driving was continued for a while and then stopped, a single flash, delivered some minutes later, would elicit a burst of responses at 3/sec. Thus, the anodal current appears to confer on the unit the capacity to retain some representation of an imposed stimulus pattern for a period of minutes.

EEG data are still apt to be presented in the form of a series of demonstrations or examples. However, computer processing of bioelectric data, demonstrated by Gerstein at this conference, shows promise of revealing certain subtle and hitherto unsuspected features of the evoked response such as prolonged after-effects.

It is no more than two decades ago that psychologists became aware of the spontaneous activity of the C.N.S. to the point of incorporating this fact into their theorizing. According to Teuber, the time has come to consider the brain as a system which operates upon its own inputs. The functional significance of sensory relay nuclei may thereby be elucidated.

Michael Reese Hospital, Chicago, Ill. J. ORBACH

A Model of the Mind. By GERALD S. BLUM. New York, John Wiley & Sons, Inc. 1961. Pp. xi, 229. $6.95.

Considering the scope of the undertaking described in this small book it is no wonder that the collaboration of eight people was required to produce it. Dr. Blum and his research team, disillusioned with what they see as the trivia of Behaviorism, have set about to fill the empty hole between S and R. They have filled it with electric wires. Their book, understandably, thus has two purposes: 1) It presents circuit diagrams, lessons for building home models of the mind, and reports of experiments intended to test how well the model will bear current, and 2) It provides shocks to a number of psychologists. In both respects, I would judge it rather successful, and recommend it as an important book for students of personality.

The general strategy of the project described in Part A of the book was to develop a comprehensive scheme for explaining everything that transpires psychologically within a person from the onset of stimulation to its reduction via effector discharge. The scheme assumes that people are integrated machines, all of whose component subsystems operate in lawful relationships to each other, and that a small number of explanatory principles should be useful to explain areas of behavior as diverse as perception and motivation, learning and logical thought, hypnosis and inhibition, psychopathology, therapy, and memory. A brief description of the evolution of schemata down to the current model, Form P, is offered together with a diagram of the latter. In essence, it consists of a seemingly electrical device in which signal inputs (stimuli) are first passed through a series of transformers, filters, and amplifiers in the "cognitive subsystem" to be then transmitted to the parallel "motoric" or "effective" subsystems from whence, after further due processing, they discharge effector devices. Between these end points, a number of concepts are invoked to explain and describe the detailed operations of the system, and these involve a terminology which blends electronic talk and psychologese: input and arousal, facilitation and feedback, inhibition and reverberation.

The second stage of the project was to *deduce* specific experiments from the model which would test its predictive power. Twelve such experiments are reported in Part B ranging from studies of associative recall and memory to inhibition of anxiety and selective blindness. Finally, Part C reverses the process of deriving ideas *from* the model and tries to incorporate some well known psychological notions *within it.* The dynamics of motivation and emotion, anxiety, frustration, inhibition, imagery, and more are all accounted for, with lengthy and virtuoso performances reserved to explain hypnosis, primary process thinking, and psychotherapy.

This book will probably be subjected to two primary kinds of criticism, 1) that the model implies a formal theory but does not actually contain one, and 2) that the experiments, all of which use hypnosis, are poor ones.

Viewing the model as the incarnation of a theory is inevitably distressing, for it dashes from Pavlov to Freud to the TV repairman with cavalier indifference to a complete portrayal of any of them. But however limited the intent of the electrical analogies it employs, they may fairly be invoked here to say that one may not object to the diagram of a light-bulb circuit on the grounds that it says little of the nature of electricity or light bulbs—it says much of their relationship. This model attempts no more than to trace the path followed "down and round" by some very complex music to see where it comes out. And if, to further mix metaphors, one may be better and more precisely led through the dark of the black box by holding (guide) wires, so much the better.

But are we better led thereby, and if so, will hypnotic experiments, each on a very few people, prove the case? Blum is keenly aware of some signal difficulties in the use of hypnosis as a research device, his experiments are ingeniously contrived to eliminate many of them, and he is, in any case, trying to demonstrate the heuristic value of his model for generating research rather than deriving proofs of any kind. The issue, however, still remains in doubt. It is no idle worship of large sample statistics that argues for more than ten subjects to be used over 12 experiments, the more so as no single study seems to employ more than six. And however painful the task of seeing each *S* for 30 or 40 hours, it is insufficient grounds for seeing too few of them. The experiments are the weakest part of the book, despite their often brilliant designs. They also sometimes seemed to me difficult to read and understand, and unfortunately sometimes lack clear summaries. This is uncharacteristic of Blum's writing, which is for the most part colorful, exciting, and lucid.

Taken in proper perspective, this book is essentially an exercise in heuristic that has two very salutary functions: 1) It encourages a needed "pretentiousness" among scientific students of human behavior by demanding and exemplifying a scientific posture towards truly complicated processes. 2) It encourages one, by its diagrammatic character, to think clearly about his own behavior schemes, demanding an accounting at every step for all his local variants of switches, relays, and wires. Were this model an entirely elegant or thoroughly predictive one, it would no longer be a model but a theory. Even taken on its own terms, holes may be discerned. Blum excoriates behaviorists, for example, for disregarding internal processes, but secretly repeats their sins by disowning consciousness. Form P has no "cognitive inputs" in its feedback system, though I would guess there may be some readers of this review who, thinking (sic) about it, will go out and buy this book. They are well advised to do so, for though Form Q will be a better model of the mind, the present one is an already good model for this branch of psychological thought.

University of Illinois PERRY LONDON

Behavioral Approaches to Accident Research. By H. H. JACOBS, *et. al.* New York, Association for the Aid of Crippled Children, 1961. Pp viii, 178. $1.50.

This paperbound volume presents the papers and discussions of a conference on the subject sponsored by the Association for the Aid of Crippled Children held at West Point, New York, October 20 and 21, 1960.

Herbert H. Jacobs of Public Service Research Institute Inc. in, "Conceptual and Methodological Problems in Accident Research," considers various definitions of accidents, questions of risk, questions of exposure measurement, sampling problems, accident classification, and mathematical models. Small probability of occurrence, sample size required, and non-normal distributions are among problems which must be met. A discussion of accident prevention research and statistical problems of evaluating it ends the paper.

Edward A. Suchman (New York City Dept. of Health) presents, "A Conceptual Analysis of the Accident Phenomenon." He considers definitions of an accident from medical, legal, common sense, and other points of view. He then considers very carefully problems of operational indexes to be used in research and their validity and shortcomings. He ends with a discussion of theoretical models, the epidemiological model and the behavioral model. Although certain individuals and different situations will have higher or lower frequency of accidents, theoretically, there can be no single causal factor, but rather a whole system of personal and environmental factors. It is pointed out that the model of "accident proneness" has not proved productive in the past.

Bernard H. Fox discusses the first two papers also on a very thoughtful level. He differentiates between "hazard," meaning objective danger, and "risk," meaning subjective estimate of hazard, and he discusses combinations of these. Drawing a parallel between epidemiological attacks on disease and accident prevention is unfortunate in that a simple source of pathology is not appropriate. A classification of immediate factors, intermediate and distant factors, is suggested.

The reviewer suggests careful reading of these three papers by anyone in accident research and especially those starting in it for the first time.

Under the title "Safety and Human Behavior," Dr. John MacIver (M.D.) of United States Steel Corp. discusses various kinds of accidents, their relation to various characteristics of individuals, notes that "increased accident susceptibility" should be substituted for the term "accident proneness" since it has not been possible to isolate such individuals with consistency. Accidents emerge from the interaction of individuals and various habits and attitudes.

In outlining "The Contribution of Experimental Psychology," James J. Gibson of Cornell University proposes a classification of the various dangers met by individuals. Accidents occur from misperception of a danger or inappropriate reaction to it.

Martin Deutsch of New York Medical College discusses adult-child relationships under "Socio-developmental Considerations in Children's Accidents." He distinguishes between mishaps and accidents, the former being a necessary part of a child's development. Underprotection may lead to more of these in the process of learning, overprotection to fewer but also to lesser development of effective methods of meeting danger.

"Cultural Factors Affecting Accidents Among Children," by Edith J. Alpenfels and Arthur B. Hays III of New York University, discusses child-adult interaction in the culture. Imitation in play, the pressure to use adult instruments, and a premium on "taking a chance" for some children and groups lead to the suggestion that there is a limit between the extremes of no accident involvement and cultural nonsurvival. Areas of cultural factors for investigation are suggested.

Alfred M. Freedman, M.D., of New York Medical College, discusses "Bio-social Approaches to Research on Childhood Accidents." A study in one area showed the most common cause of death among children was lead poisoning. A plea for research on the social matrix by the behavior scientist and concern by the community for social controls is made.

"Sociological Factors and Childhood Accidents" are discussed by Nelson N. Foote of the General Electric Company. Percentages are given of different kinds of childhood accidents which point to increased risk taking as parental protection decreases and as the child increases his development and range of exploration. Various personal and interpersonal factors are discussed and institutional factors are suggested by which society could reduce hazards.

In the first discussion, Urie Bronfenbrenner of Cornell University makes a plea for more systematic and ingenious methods of observing accident phenomena and the conditions under which they occur. Then the attempt to obtain a theoretical explanation would be more effective.

In another discussion Solomon E. Asch of Swarthmore College calls for more consideration of methods of prevention by breaking the unfavorable sequence at a key point.

Finally, Leonard S. Cottrell, Jr., Russell Sage Foundation, contributes suggestions for future reesearch on childhood accidents. John A. Clausen of the University of California, and Anatol Rapoport, University of Michigan, review and summarize the papers as a whole.

Michigan State University T. W. FORBES

Psychologist at Large: An Autobiography and Selected Essays. By Edwin G. Boring. New York, Basic Books, 1961. Pp. 371. $6.50.

Professor Boring's characteristic style—clear, frank and tolerant—illuminates this absorbing account of his life and work. The autobiography which he published in 1952 (*A History of Psychology in Autobiography,* Vol. IV. Edited by H. S. Langfeld, E. G. Boring, H. Werner, and R. M. Yerkes, 1952, Pp. 27-52) was cut from 28,000 words to 14,000. The original version is now brought to us with additions, bringing us up to 1960, through what he now considers his "best decade." The autobiography is followed by a small selection (much too small for this reader) of his letters to other psychologists: Terman, Yerkes, Koffka, Bartlett, Skinner, Sachs. Two hundred and forty-one pages are given to papers which he has selected as his most satisfactory accomplishments and a bibliography. The letters and papers are given brief introductions which place them in their setting and give his current evaluations of the ideas they contain.

Professor Boring had originally suggested the series on psychologists' autobiographies to Murchison because he had felt the need for this kind of information in writing his *History of Experimental Psychology.* When he came to write his own autobiography he tried to include the material which would interest a later historian —"endeavoring to penetrate to the springs of my own motivation and to see why my values were as they were and why they are now as they are" (p. 4). This reader feels that he is as objective about himself as anyone can be, and certainly more objective than most would be. When he was writing the 1952 version of the autobiography, Professor Boring reports, a colleague urged him not to publish it because " 'it is too frank; your graduate students will never again respect you.' I treated this advice with a proper contempt and found that respect for the maturity of graduate students wins their respect in turn" (p. 4).

The letters and papers are an essential part of the autobiography in that they sample the substance and style of his thinking. Of course, all psychologists are familiar with his historical assessments in *A History of Experimental Psychology,* and some of the papers in this collection deal with the role of the *Zeitgeist* and problems of scientific discovery. Nevertheless, it is illuminating to see these papers all together and to relate them to the events of the autobiography. The letters are especially interesting because they often show Professor Boring's thought upon topics not dealt with in formal papers. See, for example, the letter to Lewis Terman (pp. 101-104) on the international situation in 1947, and what psychologists should or should not do in support of the cause of peace.

I have mentioned above that the selection of letters seems all too brief. Having exchanged some letters with Professor Boring on Emmert's Law and size-constancy while I was still an instructor, and knowing the pains he took in discussing a point of difference with someone he scarcely knew, I suspect that a sampling of the correspondence with students and junior colleagues might have been extremely rewarding, and would probably have revealed still other aspects of his thinking.

In addition to Professor Boring's individual development and thought, the book also provides interesting glimpses of Titchener and the Cornell Department of his day, and of the development of the present Department of Psychology at Harvard. To those of us who never knew Titchener, his influence upon his students seems mysterious, and only partly explained by his erudition and his psychological system.

Perhaps it will always seem a puzzle to us, but Professor Boring has succeeded in at least partially explaining the phenomenon by describing the ways in which it affected him. He has the advantage that he accepted the domination with reservations and considered himself intellectually a rebel within the group. As for the development of the Harvard Laboratory, we are given the chronology, but Professor Boring modestly credits the *Zeitgeist* with the main responsibility for the growth and emancipation of the Department from Philosophy. Here, we may suspect, there is some failure of objectivity, and it will be left to others to describe his contributions more specifically.

One good measure of the success of this book is that we can only wish that it were longer. Another is that it achieves its purpose since most readers will finish the book with the conviction that they understand how Professor Boring developed into the eminent scientist known to all psychologists today. All who study *A History of Experimental Psychology* (and who doesn't?) should get acquainted with the man who produced it.

T. A. R.

Experiments in Hearing. By GEORGE VON BÉKÉSY. New York, McGraw-Hill, 1960. Pp. x, 745. $25.00.

In this volume Von Békésy brings together his experimental work in audition which has spanned a little more than three decades. Except for the four introductory chapters the book is constituted of material previously published in professional journals. Normally one would not expect the republication of experimental work in its original form to be worthwhile, yet in this instance the effort clearly has merit. Because much of Von Békésy's work first appeared in German and in journals with limited circulation in this country, the appearance of *Experiments in Hearing* facilitates immeasurably access to the author's important work.

The first two chapters serve as a general introduction to the topic of audition. Many young experimenters—and perhaps some old ones—would find it profitable to read these chapters, for in them Von Békésy intimates his general philosophy as a scientist. The third and fourth chapters complete the introductory section of the book and treat generally the experimental techniques so essential to an understanding and an evaluation of what follows.

Although the compilation of his papers is of special interest to sensory psychologists, it probably is not of interest to most psychologists because of its highly specialized nature. Nevertheless, for those working in audition it serves as an excellent source book.

In certain respects the title of the volume is misleading. Little attention is focused upon psychophysical matters. In fact, only about one-third of the volume treats material classically considered in the area of hearing. Most of the volume deals with the physics of sound, the physiology of the ear, and cochlear mechanics. Of special significance, however, is the detailed treatment of his experimental methodology. One is seldom in doubt about what Von Békésy did or how he did it.

The author makes a convincing case for the use of models in the study of the ear and his treatment of frequency analysis in the cochlea is excellent. Moreover, his comparison of auditory and cutaneous sensibilities is clearly and interestingly presented. Unfortunately, he has not yet made plain the nature of his traveling wave, whether it be the result of energy transmission through the membrane or successive phase lags in hydraulic pressure in the fluids bounding the basilar membrane.

Professor E. G. Wever is responsible for translating the German papers and editing the entire volume. The significance of his contribution ought not to go unmentioned.

Finally, Von Békésy's genius as an experimentalist cannot escape the careful reader, and no serious student of audition can ignore the tremendous contributions of the author set down in this volume.

University of Delaware W. L. GULICK

Thought and Language. By LEV SEMENOVICH VYGOTSKY. Edited and translated by Eugenia Hanfmann and Gertrude Vakar. Introduction by J. S. Bruner and Comments by Jean Piaget. New York, M.I.T. Press and John Wiley and Sons, 1962. Pp. xxi, 168. $4.95.

Until recently Vygotsky (1896-1934) has been little more than a name in the English-reading world. There is a fragment, for instance, in Saporta's *Psycholinguistics* (1961), and Hanfmann and Kasanin have paid tribute to him in their studies of schizophrenic thinking. For the most part, however, western psychologists were unaware of the fact that in the 1920's and 1930's Russia was rearing a psycholinguist who will rank in history with the best. *Thought and Language* is an edited translation of a number of manuscripts which Vygotsky had assembled shortly before his untimely death at the age of 38 yr. Hanfmann and Vakar have done a magnificent job with what must have been a somewhat muddled set of papers. It is to be hoped that they will give us more from the same source. Bruner's graceful introduction brings Vygotsky to life for us. Piaget's equally graceful comments, included in a supplement, express his regret that he had not known Vygotsky earlier.

Now for a word of criticism, which may sound brutal. Vygotsky's book is good, but not great. It should be read by all students of the psychology of language and thinking, simply because it is provocative and because we should read some of the post-Pavlovian Russians. Vygotsky was not exactly in step with the Party. It is disturbing, however, to find a brilliant writer who is so massively ignorant of both of the history of psychology and of the history of linguistics. Let us say that he was writing in the late 1920's. By that time he should have known that the battle against elementarism had been pretty well won by the Gestaltists. Long before that the Young Grammarians had been defeated, and the new movement in linguistics had begun. It is probable that Vygotsky could read English, German and French; for he quotes liberally, especially from Stern and Piaget. He has a reference to the Von Humboldts, but he fails to distinguish between the two brothers; and he seems not to have remembered Herder, who launched the whole psycholinguistic movement.

In this reviewer's opinion, Vygotsky has something important to say. In almost utter innocence of what was being done in other countries he set up ingenious experiments on concept attainment in children. The fact that his findings are neither new nor shocking should not diminish our appreciation.

Cornell University ROBERT B. MACLEOD

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages.
Listing here does not, however, preclude their later review.)

ALEXANDER, THERON, *Psychotherapy in Our Society.* Englewood Cliffs, N.J., Prentice-Hall, 1963, Pp. ix, 181. Cloth, $3.95; Paper, $1.95.

BARKER, ROGER G., *The Stream of Behavior.* New York, Appleton-Century-Crofts, 1963. Pp. x, 352. $6.75.

BAUMRIN, BERNARD (ed.), *Philosophy of Science.* Vol. 1, 1961-62, *The Delaware Seminar.* New York, John Wiley, 1963. Pp. xvi, 370, $9.75.

BECKER, HOWARD S., *Outsiders: Studies in the Sociology of Deviance.* New York, Free Press of Glencoe, 1963. Pp. x, 179. $5.00.

BENNETT, EDWARD; DEGAN, JAMES, and SPIEGEL, JOSEPH (eds.), *Human Factors in Technology.* New York, McGraw-Hill, 1963. Pp. xiii, 685. $14.50.

BERELSON, BERNARD (ed.), *The Behavioral Sciences Today.* New York, Basic Books, 1963, Pp. viii, 278. $4.95.

BERKOWITZ, LEONARD, *Aggression: A Social Psychological Analysis.* New York, McGraw-Hill, 1962. Pp. xv, 361. $7.95.

BION, W. R., *Learning from Experience.* New York, Basic Books, 1963. Pp. xii, 111. $4.00.

BOCKOVEN, J. S., *Moral Treatment in American Psychiatry.* New York, Springer Publ. Co., 1963. Pp. ix, 116. $3.00 (paperbacks).

BOSS, MEDARD, *Psychoanalysis and Daseinsanalysis.* New York, Basic Books, 1963. Pp. viii, 295. $8.50.

BRESCIA, JOSEPH, *Journey into Superconsciousness.* New York, Vantage Press, 1963. Pp. 118. $2.95.

BUNGE, MARIO, *Causality. The Place of the Causal Principle in Modern Science.* Cleveland, Ohio, World Publ. Co., 1963. Pp. xx, 394, $2.25 (paperback).

BUTLER, JOHN M.; RICE, LAURA N., and WAGSTAFF, ALICE K., *Quantitative Naturalistic Research.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. v, 122. $4.95.

CHOISY, MARYSE, *Sigmund Freud: A New Appraisal.* New York, The Citadel Press, 1963. Pp. 141. $1.75 (paperback).

CLARK, JAMES V., *Education for the Use of Behavioral Science.* Los Angeles, Univ. of Calif. Institute of Industrial Relations, 1962. Pp. 91. $2.25 (paperback).

COFER, CHARLES N., and MUSGRAVE, BARBARA S. (eds.), *Verbal Behavior and Learning.* New York, McGraw-Hill, 1963. Pp. x, 397. $8.50.

COMMITTEE ON HOMOSEXUAL OFFENSES AND PROSTITUTION (with a foreword by KARL MENNINGER), *The Wolfenden Report.* New York, S. D. Stein and Day, 1963. Pp. 243. $5.95.

DE CECCO, JOHN P., *Human Learning in the School. Readings in Educational Psychology.* New York, Holt, Rinehart and Winston, 1963. Pp. xix, 636. (paperback).

DENNIS, WAYNE, *Readings in Child Psychology.* 2nd ed. Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. xi, 532. Text, $6.00; Trade, $8.00.

DOBRINER, WILLIAM M., *Class in Suburbia.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. xii, 166. Cloth, $4.95; Paper, $1.95.

ETHRIDGE, JAMES M. (ed.), *Contemporary Authors: An International Bio-Bibliographical Guide to Current Authors and Their Works,* Vol. 7. Detroit, Mich., Gale Research Co., The Book Tower, 1962. Pp. 245. Annual subscription (four volumes), $28.50; Individual volumes, $10.00.

FARBER, SEYMOUR M., and WILSON, ROGER H. L., (eds.), *Conflict and Creativity. Control of the Mind,* Part 2. New York, McGraw-Hill, 1963. Pp. xiii, 360. Hardcover, $6.50; Paperback, $2.95.

FESTINGER, LEON; SCHACHTER, STANLEY, and BACK, KURT, *Social Pressures in Informal Groups. A Study of Human Factors in Housing.* Stanford, Calif., Stanford

Univ. Press, 1963. Pp. x, 197. $3.00. (Reissue; first publ. by Harper & Bros. in 1950).

FLAVELL, JOHN H., *The Developmental Psychology of Jean Piaget.* New York, D. Van Nostrand, 1963. Pp. xvi, 472. $8.75.

FOGEL, LAWRENCE J., *Biotechnology: Concepts and Applications.* New York, Prentice-Hall, 1963. Pp. xvii, 826. $22.00.

FRAISSE, PAUL. *The Psychology of Time.* New York, Harper and Row, 1963. Pp. vi, 343. $5.50.

FRAISSE, PAUL, and PIAGET, JEAN, *Traité de Psychologie Expérimentale. I. Histoire et Méthode; II. Sensation et Motricité; VI. La Perception.* Paris, Presses Universitaires de France, 1963. Pp. 189; 159; 229. F. 12; F. 12; F. 15.

FREUD, SIGMUND, *Three Essays on the Theory of Sexuality* (translated and revised by JAMES STRACHEY). New York, Basic Books, 1963. Pp. xviii, 130. $3.95.

FRY, EDWARD, *Teaching Machines and Programmed Instruction.* New York, McGraw-Hill, 1963. Pp. viii, 244. $5.95.

GOOD, I. J. (ed.), *The Scientist Speculates. An Anthology of Partly-Baked Ideas.* New York, Basic Books, 1963. Pp. xvii, 413. $6.95.

GOUDE, GUNNAR, *On Fundamental Measurement in Psychology.* Stockholm, Almqvist and Wiksell, 1962. Pp. 176. Price, Sw.kr. 18:—.

GREENBERG, JOSEPH (ed.), *Universals of Language.* Cambridge, Mass., MIT Press, 1963. Pp. x, 269. $6.00.

HARDING, M. ESTHER, *Psychic Energy. Its Source and Its Transformation,* 2nd ed. New York, Pantheon Books, 1963. Pp. xix, 497. $6.00.

HILL, WINFRED F., *Learning: A Survey of Psychological Interpretations,* San Francisco, Calif., Chandler Publ. Co., 1963. Pp. x, 227. $1.95 (paperback).

HOLLAND, R. F. (ed.), *SENSATIONALISM AND SCIENTIFIC EXPLANATION.* New York, Humanities Press, 1963. Pp. v, 149. $3.25.

HUNT, EARL B., *Concept Learning: An Information Processing Problem.* New York, John Wiley, 1962. Pp. viii, 286, $7.50.

JACOBI, JOLANDE, *The Psychology of C. G. Jung,* new rev. ed., New Haven, Conn., Yale Univ. Press, 1963. Pp. xiv, 193. Cloth, $6.00; Paper, $1.75.

JERSILD, ARTHUR T., *The Psychology of Adolescence,* 2nd ed. New York, Macmillan Co., 1963. Pp. xi, 468. $6.00.

JOHNSON, WENDELL; DARLEY, FREDERIC L., and SPRIESTERBACH, D. C., *Diagnostic Methods in Speech Pathology.* New York, Harper and Row, 1963. Pp. xv, 347. $7.50.

JOURARD, SIDNEY M., *Personal Adjustment. An Approach through the Study of Healthy Personality,* 2nd ed. New York, Macmillan Co., 1963. Pp. xiii, 477. $6.50.

JUNG, C. G., *Memories, Dreams, Reflections* (edited and recorded by ANIELA JAFFE), New York, Pantheon Books, 1963. Pp. xvii, 398. $7.50 .

KARDINER, ABRAM, and PREBLE, EDWARD, *They Studied Man.* New York, World Publ. Co. (Mentor Book), 1963. Pp. viii, 255. $.75 (paperback).

KENDLER, HOWARD H., *Basic Psychology.* New York, Appleton-Century-Crofts, 1963. Pp. xiii, 750. $7.50.

KENNY, ANTHONY (edited by R. F. HOLLAND), *Action, Emotion and Will.* New York, Humanities Press, 1963. Pp. vii, 245. $5.00.

KOCH, SIGMUND, *Psychology: A Study of a Science. Study II. Empirical Substructure and Relations with Other Sciences.* Vol. 5, 6. New York, McGraw-Hill, 1963. Pp. xii, 967; xii, 791. $13.50; $12.50.

KOLLE, KURT, *Introduction to Psychiatry.* New York, The Citadel Press, 1963. Pp. 71, $1.50 (paper).

LAZARUS, RICHARD S., *Personality and Adjustment.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. x, 118. Cloth $3.95, paper $1.50.

LEVINE, LOUIS S., *Personal and Social Development. The Psychology of Effective Behavior.* New York, Holt, Rinehart and Winston, 1963. Pp. xii, 510. $6.50.

LEVINSON, ELIZABETH J., *Retarded Children in Maine: A Survey and Analysis.* Orono, Univ. of Maine Press, 1962. Pp. xii, 265. $2.00 (paperback).

LEVITAS, G. B. (ed.), *The World of Psychology. Vol. I. The Worlds of Perception,* *Man and His Emotions; Vol. II, Identity and Motivation.* New York, George Braziller, 1963. Pp. xi, 563; vi, 583. $17.50.

Levy, Leon H., *Psychological Interpretation.* New York, Holt, Rinehart and Winston, 1963. Pp. xii, 368. $6.50.

Lewis, E. Vernon, *Statistical Analysis. Ideas and Methods.* Princeton, N.J., D. Van Nostrand Co., 1963. Pp. x, 484. $8.00.

Lyman, Howard B., *Test Scores and What They Mean.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. xv, 223. Text, $3.75; Trade, $5.00.

Maslow, Abraham H., *Toward a Psychology of Being.* Princeton, N.J., D. Van Nostrand, 1962. Pp. ix, 214. $1.95 (paperback).

Mazlish, Bruce (ed.), *PSYCHOANALYSIS AND HISTORY.* Englewood Cliffs, N. J., Prentice-Hall, 1963. Pp. vi, 183. $2.25 (paperback).

McKenna, A. E., *Les Reactions Psycho-Physiologiques à la Douleur.* Paris, Beatric-Nauwelaerts, 1963. Pp. 126. Price, FB 120.

Messick, Samuel, and Ross, John (eds.), *Measurement in Personality and Cognition.* New York, John Wiley and Sons, 1962. Pp. xi, 334. $5.75.

Miller, George A., *Psychology. The Science of Mental Life.* New York, Harper and Bros., 1962. Pp. xi, 388. $4.95 (text ed.).

Moser, Leslie E., and Moser, Ruth Small, *Counseling and Guidance: An Exploration.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. xii, 432. $6.95.

Mussen, Paul Henry; Conger, John Janeway, and Kagan, Jerome, *Child Development and Personality,* 2nd ed. New York and Evanston, Harper and Row, 1963. Pp. xiv, 625. $7.25.

Nelson, Marie Coleman, *Paradigmatic Approaches to Psychoanalysis: Four Papers.* New York, Dept. of Psychology, Stuyvestant Polyclinic, 1962. Pp. 68. $2.00 (paperback).

Palermo, David S., and Lipsitt, Lewis P. (eds.), *Research Readings in Child Psychology.* New York, Holt, Rinehart and Winston, 1963. Pp. xiii, 572. $7.50.

Pande, N. R. W., *Abilities and Education.* Maharashtra, India, Shri Mouni Vidyapeeth, 1962.

Pandey, Ram Shakal, *Our Adolescents: Their Interests and Education.* Agra, India, Lakshmi Narain Agarwal, 1963. Pp. xx, 387. 15 rupees.

Pearce, Jane, and Newton, Saul, *The Conditions of Mental Growth.* New York, The Citadel Press, 1963. Pp. 444. $6.95.

Peatman, John G., *Introduction to Applied Statistics. For Students of Psychology and the Behavioral Sciences.* New York, Harper and Row, 1963. Pp. xv, 458. $7.95.

Pritchard, D. G., *Education and the Handicapped. 1760-1960.* New York, Humanities Press, 1963. Pp. viii, 250, 12. $5.50.

Rao, M. S., *An Introduction to a New Psychiatry through a New Psychopathology or Abnormal (Clinical) Psychology.* Subbash Marg, Jaipur, India, Rao's Institute of Psychology, 1962. Pp. xv, 514. $3.00.

Reik, Theodor, *The Need to be Loved.* New York, Farrar, Strauss, 1963. Pp. 276. $4.95.

Reik, Theodor, *Of Love and Lust.* New York, Farrar, Straus, 1963. Pp. xiv, 623. $2.95 (paperback). (Reprint; first published 1941).

Rethlingshafer, Dorothy, *Motivation as Related to Personality.* New York, McGraw-Hill, 1963. Pp. ix, 388. $7.95.

Riley, John W., Jr. (ed.), *The Corporation and its Public. Essays on the Corporate Image.* New York, Wiley and Sons, 1963. Pp. xii, 195. $5.50.

Rosca, Alexendru (ed.), *Tratat de Psihologie Experimentala.* Bucharest, Roumania, Editura Academiei Republicii Populare Romine, 1963. Pp. 650, plus Abstract in English, 30 pp.

Rugg, Harold, *Imagination. An Inquiry into the Sources and Conditions that Stimulate Creativity.* New York, Harper and Row, 1963. Pp. xxii, 361. $6.95.

Schramm, Wilbur (ed.), *The Science of Human Communication.* New York, Basic Books, 1963. Pp. vii, 158. $4.50.

Seidman, Jerome M. (ed.), *Educating for Mental Health. A Book of Readings.* New York, Thomas Y. Crowell Co., 1963. Pp. xi, 580. $4.50 (paperback).

SELLS, S. B., and SELLS, HELEN F., (eds.), *Bioelectronics Directory.* 2nd ed. Fort Worth, Texas, Texas Christian Univ., 1963. Pp. ii, 64.

SILBERMANN, ALPHONS (translated by CORBET STEWART), *The Sociology of Music,* New York, Humanities Press, 1963. Pp. v, 216. $5.50.

SIMON, BRIAN and SIMON, JOAN, *Educational Psychology in the U.S.S.R.* Stanford, Calif., Stanford Univ. Press, 1963. Pp. xi, 283. $6.75.

STEINER, GARY A., *The People Look at Television. A Study of Audience Attitudes.* New York, Knopf, 1963. Pp. xvii, 422. $7.95.

STEVENSON, HAROLD W.; KAGAN, JEROME, and SPIKER, CHARLES (eds.), *Child Psychology. The Sixty-Second Yearbook of the National Society for the Study of Education, Part I.* Chicago, Ill., Univ. of Chicago Press, 1963. Pp. x, 550, vi. $6.50.

SWARTZ, PAUL, *Psychology: The Study of Behavior.* Princeton, N.J., D. Van Nostrand, 1963. Pp. xix, 451. $6.95.

TAYLOR, CALVIN W., and BARRON, FRANK (eds.), *Scientific Creativity: Its Recognition and Development.* New York, Wiley and Sons, 1963. Pp. xxiv, 419. $7.95.

TOWNSEND, EDWARD ARTHUR, and BURKE, PAUL J., *Statistics for the Classroom Teacher. A Self-Teaching Unit.* New York, Macmillan, 1963. Pp. 76. $1.95 (paperback).

TYLER, LEONA E., *Tests and Measurements.* Englewood Cliffs, N.J., Prentice-Hall, 1963. Pp. xi, 116. Cloth $3.95, Paper, $1.50.

VENESS, THELMA, *School Leaves. Their Aspirations and Expectations.* New York, Humanities Press, 1963. Pp. xxv, 252. $5.00.

WATSON, ROBERT I., *The Great Psychologists. Aristotle to Freud.* Philadelphia, Pa., J. B. Lippincott, 1963. Pp. xiii, 572. $5.50.

WATTS, ALAN W., *Psychotherapy East and West.* New York, The New American Library of World Literature, 1963. Pp. x, 160. $.60 (paperback).

WEST, D. J., *Psychical Research Today.* Baltimore, Md., Penguin Books, 1963. Pp. 224. $.85 (paperback).

WHITING, BEATRICE B. (ed.), *Six Cultures. Studies of Child Rearing.* New York, John Wiley, 1963. Pp. vi, 1017. $12.50 .

WIEMAN, HENRY NEILSON, *Man's Ultimate Commitment.* Carbondale, Ill., Southern Illinois Univ. Press, 1963. Pp. x, 318. Paper, $1.95; Cloth, $6.00.

THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. STANLEY HALL

VOL. LXXVI DECEMBER, 1963 NO. 4

A TEST OF THE INVARIANCE OF THE RATIO OF PERCEIVED SIZE TO PERCEIVED DISTANCE

By WALTER C. GOGEL, EUGENE R. WIST, and GEORGE S. HARKER,
Fort Knox, Kentucky

Perceived size and perceived distance are often considered to be related. An expression of this relation is given by the hypothesis of the invariance of the ratio of perceived size to perceived distance, which states that the perceived absolute size (S') of an object of constant visual angle (θ) is proportional to its perceived absolute distance (D'). In equational form the hypothesis may be expressed as:

$$S' = K\,\theta\,D', \qquad [1]$$

where K is a constant.[1] It has been pointed out that Equation [1] parallels the physical equation:

$$S = \theta\,D, \qquad [2]$$

where S and D are physical size and distance respectively and θ is reasonably small such that an $\theta = \theta$ in radians.[2] The formal similarity of Equations [1] and [2] suggests a parallel similarity in the definitions of physical extents and perceived extents. The physical distance D is the distance from the eye to the object. The perceived distance D' is the O's perception of the distance of the object from himself. The perceived distance D' is termed perceived absolute distance to distinguish it from a perception of the depth

---

* Received for publication August 28, 1962. This research was conducted at the U.S. Army Medical Research Laboratory, Fort Knox, Kentucky. The authors wish to thank Don L. Carmichael, Harlyn W. Lacey, Harold V. Kulman, and Shuna A. Miah for their assistance in collecting and analyzing the data of this study. W. C. Gogel is now at the Civil Aeromedical Research Institute, Oklahoma City, Oklahoma.

[1] F. P. Kilpatrick and W. H. Ittelson, The size-distance invariance hypothesis, *Psychol. Rev.* 60, 1953, 223-224.

[2] Kilpatrick and Ittelson, *op. cit.*, 223-224.

between two objects, *i.e.* from a perception of a relative depth. The physical size, S, is the physical height (or width) of the object without regard to the height (or width) of surrounding objects. Similarly, S' is the *O*'s perception of the height or width of the object without regard to the perceived height or width of surrounding objects. The term S' is called perceived absolute size to distinguish it from a perception of the relative size of objects, *i.e.* from a perception of the ratio of the size of one object with respect to another. The hypothesis of the invariance of the ratio of perceived size to perceived distance is a statement that one perceptual absolute determines another.

The evidence for the validity of this hypothesis is equivocal.[3] Two methods of testing it can be differentiated.[4] (1) One method is to correlate the size and distance judgments from different *O*s for an object of constant physical size at a constant physical distance. The question answered by this method is whether the differences in the perceptions of absolute size from different *O*s are related to the corresponding differences in the perceptions of absolute distance. This method usually has not yielded correlations in support of the hypothesis of invariance.[5] (2) The second method is to determine the correlation between the perceived size and the perceived distance of objects as a function of distance rather than of *O*s. A basic observation upon which the hypothesis of invariance rests is that an object subtending a constant visual angle seems to grow larger as it appears to recede into the distance. This type of observation implies the second method. This method, by providing a potentially large range of differences in perceived size and perceived distance, can constitute a sensitive test of the hypothesis of invariance. Evidence in support of a positive relation between perceived size and perceived distance can be found in studies of size-constancy. Indeed, it is often found that under conditions in which the perceived distance of an object of constant physical size obviously increases with increasing physical distance, the perceived size of the object remains essentially unchanged (perfect size-constancy). It follows that, under the same conditions, the perceived size of an object of constant visual angle would increase with distance. Clearly, under these circumstances, perceived size and perceived distance are correlated.

The perceived size measured in the usual size-constancy study is perceived relative size, *i.e.* one visual size is judged with respect to another. For constant viewing conditions, however, it is reasonable to expect that similar (proportional) results would be obtained from a method which measured perceived absolute size. Thus, it follows that size-constancy provides a demonstration that S' and D' can be related positively. The issue in the present study, therefore, is not whether perceived absolute size and perceived absolute distance in a visually rich environment are related, for an object of constant θ and varying D. A positive correlation under these conditions is expected. Instead, the question is whether the relation obtained between S' and D' is linear and unchanging. If Equation [1] is valid, K should be completely independent of the conditions of observation. For a constant θ and varying D, Equation [1] should predict S' in all situations in which D' can be demonstrated to occur, regardless of the manner in which D' was produced.

---

[3] Kilpatrick and Ittelson, *op. cit.*, 223-231; William Epstein, John Park, and Albert Casey, The current status of the size-distance hypothesis, *Psychol. Bull.*, 58, 1961, 491-514.

[4] E. E. Rump, The relationship between perceived size and perceived distance, *Brit. J. Psychol.*, 52, 1961, 113-114.

[5] H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-426; Noël Jenkins and Ray Hyman, Attitude and distance-estimation as variables in size-matching, this JOURNAL, 72, 1959, 68-76; V. R. Carlson, Overestimation in size-constancy judgments, this JOURNAL, 73, 1960, 199-213; Rump, *op. cit.*, 116.

The present study tests the hypothesis of the invariance of the ratio of perceived size and distance in a situation in which the cues for perceived depth were systematically modified. The modification of the cues for depth was accomplished by using optical devices which produced magnification (or minification) of the effective interpupillary distance (base) of the eyes. Magnification of the base, by itself, results in essentially the magnification of only the binocular disparities in the field of view. Equation [1], if valid, must apply to these as well as to more normal viewing conditions.

This study was also concerned with the problem of the general relation between retinal distortions and changes in the resulting perceptions. Ames has demonstrated with the use of aniseikonic lenses that the perceptual consequences of distorting the optical image can be different depending upon the type of environment which is being viewed.[6] This is probably the result of the interaction of cues, only some of which were affected by the aniseikonic lenses. In the present study, only the binocular disparity was modified by magnification or minification of the base. It has been demonstrated previously that the perception of depth from binocular disparity and the perception of frontal size (the perceived size of objects presented parallel to the frontal plane of O) are interrelated.[7] Thus, it is possible that modifications of cues to the depth-dimension can result in modifications in the perception of the size of an object presented in a frontoparallel plane even though the visual angle of the object is unchanged.

METHOD AND PROCEDURE

(1) *Apparatus: (a) Interpupillary instruments.* The modification of the binocular disparities between objects in the field of view by magnification or minification of

---

[6] Adelbert Ames, Jr., Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments, this JOURNAL, 59, 1946, 333-357.

[7] W. C. Gogel, The perception of space with binocular disparity cues, USAMRL Report No. 379, Ft. Knox, Ky., 1959, 1-24; Perceived frontal size as a determiner of perceived binocular depth. *J. Psychol.,* 50, 1960, 119-131; The perception of shape from binocular disparity cues, *idem,* 179-192; The perception of a depth interval with binocular disparity cues, idem, 257-269.

the effective interpupillary base is illustrated in Fig. 1. In this figure, the binocular disparity between Objects *e* and *f* is $\alpha_e - \alpha_f$ while the distance from *O* to Objects *e* and *f* is D_e and D_f respectively, with X_{ef} being the distance between the objects. The lateral distance between the eyes of *O* (the interpupillary distance) is given as '*A*.' It is readily shown that for Fig. 1b,

$$\alpha_e - \alpha_f = AX_{ef}/(D_e^2 + D_eX_{ef}) \qquad [3]$$

where $\alpha_e - \alpha_f$ is expressed in radians. The modification of the effective interpupillary distance '*A*' by the use of mirrors is diagrammed in the same figure. Fig. 1a illustrates the situation in which the effective interpupillary distance is twice that shown in Fig. 1b. In Fig. 1a, *O* sees Objects *e* and *f* with the right eye by means of mirrors M_2 and M_4 and with the left eye by means of M_1 and M_3. From Equation [3] $\alpha_e - \alpha_f$ in Fig. 1a is twice that in Fig. 1b. Similarly, in Fig 1c, the effective interpupillary distance is one-half that of Fig. 1b and the binocular disparity between Objects *e* and *f* is also halved. Figs. 1a and 1c are examples of what are called, respectively, base-magnification and base-minification. An increase or decrease in the effective distance between the eyes by the use of such mirror devices produces an increase or decrease, respectively, in the binocular disparities between objects in the field of view.[8]

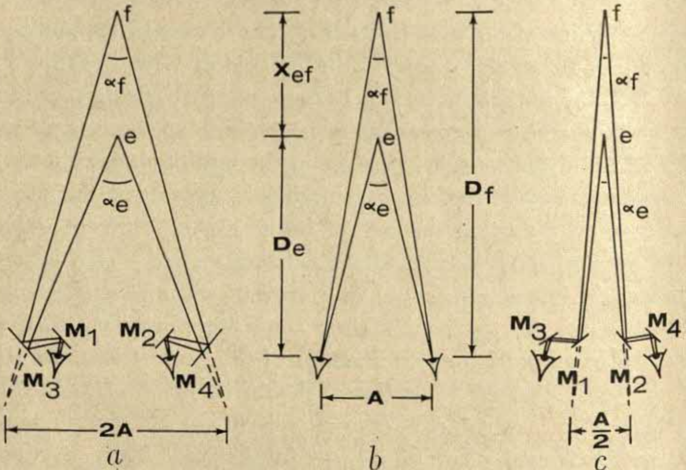


FIG. 1. SCHEMATIC REPRESENTATION OF THE THREE DIFFERENT BASES

The length of the optical path from Objects *e* and *f* to the eyes is slightly greater in Figs. 1a and 1c than in Fig. 1b. Thus, the retinal size of the frontoparallel component of Objects *e* and *f* is slightly less in the cases of Figs. 1a and 1c than in the case of Fig. 1b. When, however, D_e is large compared with the effective interpupillary distance (as is usually the case), this difference is negligible. Thus, a change in the effective interpupillary distance modifies the binocular disparities between objects without appreciably modifying the retinal size of their frontoparallel components.

---

[8] J. von Kries, in Herman von Helmholtz, *Treatise on Physiological Optics,* Vol. III (J.P.C. Southall, ed.) 1924, 660-661.

In the present study, mirrors M_3 and M_4 were adjusted so that the convergence of the O's eyes for an object at 381.0 cm. (12.5 ft.) was constant regardless of the value of base used. Since the average distance of the targets was 381.0 cm. from *O*, this minimized the differences between the average convergence value of the eyes as a function of using the different bases. Thus, the disruption of the normal accommodation-convergence relation for the different interpupillary distances was reduced.

Each *O* viewed a visual alley using three different interpupillary distances. An increased and a decreased interpupillary distance were produced by devices using front surface mirrors and the principles illustrated in Figs. 1a and 1c, respectively. The third interpupillary distance was the normal interpupillary distance of *O* as illustrated in Fig. 1b. The increased, normal, and decreased interpupillary distances will be referred to as increased, normal, and decreased bases, respectively.

Each instrument had an attached mask containing apertures of a size such that *O* could see the alley only by means of the appropriate mirrors. The mask was translucent to minimize *O*'s awareness of the apertures. With the normal base, the mirrors were removed but the mask with its restrictive apertures remained. A head- and chin-rest located at the viewing position immobilized *O*'s head during the judgments and supported the interpupillary devices. The average increased, normal, and decreased bases were 12.4, 6.5, and 3.2 cm., respectively. Each base provided a field of view approximately 24° wide and 13° high.

(b) The alley. The visual alley was 1043 cm. (34.2 ft.) long and 128.0 cm. (4.2 ft.) wide. The walls and floor were covered with tan cloth. The left wall was rigid while the right wall consisted of curtains. The floor of the alley was formed by a series of table tops located 26.0 cm. (10.2 in.) below the eyes of *O*. The folds in the curtains, the texture of the cloth, and the edges of the tables provided adequate distance cues so that the alley appeared to *O* to extend in distance. The alley was illuminated by a double row of 8-ft. fluorescent tubes extending the length of the alley and mounted on the ceiling 5 ft. above the alley floor.

Black rectangular targets, whose size and distance were to be judged, were presented in a frontoparallel plane on a stand at either 152.4, 304.8, 457.2, or 609.6 cm. (5, 10, 15, or 20 ft.) from *O*. The rectangles were 7.6, 15.2, 22.9, and 30.5 cm. (3, 6, 9, and 12 in.) wide at the respective distances, so as to subtend a constant visual angle (θ) of 2°52′. The width of a rectangle was always three times as great as its height and its lower edge was always at the height of *O*'s eyes. Only one rectangle was presented at a time.

(2) Measuring techniques. Measurements of D' and S' required that some modality other than vision be involved in indicating the extent (size or distance) which was perceived visually. If a visual indicator were used in this process a visual comparison between the indicator and the extent being measured would occur and a perception of relative rather than absolute extent would result. For this reason, kinesthesis was used for the measurement of perceived absolute size and distance.

(a) Measurement of perceived absolute distance. The kinesthetic method used in the measurement of perceived absolute distance consisted of a throwing technique. Parallel to the illuminated visual alley but separated from it by curtains was

a dark, throwing alley. The distance at which *O* perceived the black rectangle to be from himself (perceived absolute distance) was measured by having *O* throw darts down the throwing alley to land at the apparent distance of the rectangle. The floor of the throwing alley was at the height of the bottom of the rectangle and consisted of fiberboard into which the darts could readily penetrate. *O* looked through the interpupillary device and, while looking at the rectangle in the visual alley, threw a dart as far down the throwing alley as the rectangle appeared to be from himself. The curtain prevented *O* from seeing the results of his throwing. At no time did any dart land in the visual alley.

The throwing technique has been demonstrated to be a sensitive indicator of the perceived distance of an object and gives results which, while not necessarily identical to perceived absolute distance, are at least linearly related to it.[9]

(b) Measurement of perceived size. The measurement of perceived absolute size at the four distances in the visual alley proceeded in two steps. The first step consisted of an absolute size-judgment. The apparatus used for making this judgment was located at the position of *O*. It contained a pair of aluminum rods invisible to *O* which were vertically positioned at about the level of his knees. *O* could grasp these rods, one in each hand, and could move the right rod laterally with respect to the stationary left rod. This device was called the hand adjustable apparatus. Two additional rods were also used. These were black rods (0.6 cm. in diameter) and were visible to *O*. These rods were vertically presented at a distance of 152.4 cm. from *O* in the visual alley. The right black rod was moveable laterally by the experimenter (*E*) with respect to the left black rod. A small (2.5 cm. high) black barrier extending across the visual alley at the base of the black rods prevented *O* from seeing the slot in which *E* could move the right rod. For the absolute size-judgments, the lateral distance between the black rods was adjusted by *E* to a fixed value of 30.5 cm.

To indicate perceived absolute size, *O* grasped the two aluminum rods of the hand adjustable apparatus and moved the right rod laterally with respect to the stationary left rod until the lateral distance between the centers of the rods seemed to be the same as the 30.5 cm. distance between the black rods in the visual alley. A centimeter scale attached to the moveable rod of the apparatus enabled *E* to measure the magnitude of the hand-adjustment. The hand-adjustment measured the absolute rather than the relative size perceived since the process of measurement did not involve the comparison of two visual sizes. This method of measuring perceived absolute size has been used in previous studies.[10]

The second step in the measurement of perceived size consisted of a series of relative size-adjustments in which the lateral distance between the two black rods at 152.4 cm. in the visual alley was adjusted to apparently duplicate the width of the simultaneously visible rectangle (of 2°52′ angular size) at either 152.4, 304.8, 457.2, or 609.6 cm. A centimeter scale invisible to *O* permitted *E* to measure the adjusted distance between the black rods. Since the visual distance between the black rods was adjusted to equal perceptually the visual width of the black rectangle (a comparison of two visual widths) a judgment of relative not absolute size was made.

---

[9] W. C. Gogel, B. O. Hartman, and G. S. Harker, The retinal size of a familiar object as a determiner of apparent distance, *Psychol. Monogr.*, 71, 1957, (No. 442), 1-16; Gogel, *op. cit.*, 1960, 119-131.

[10] Gogel, Convergence as a determiner of perceived absolute size, *J. Psychol.*, 53, 1962, 91-104; The effect of convergence on perceived size and distance, *idem*, 475-489.

The measures of the apparent absolute (frontal) size at one distance, and the apparent relative (frontal) size at a number of distances, permitted *E* to calculate the apparent absolute (frontal) sizes at each of the distances. For example, if the object at the 152.4 cm. distance appeared to be 30 cm. wide, any other object at any other distance which appeared to be twice as wide as the first object must have appeared to be 60 cm. wide. Thus, the perceived absolute sizes of the rectangles at the respective distances could be determined by using the proper combination of the two types of size-measures.

(3) Observers. Twelve right handed men were used as *O*s in the main part of this study. These *O*s had an uncorrected visual acuity in each eye of at least 20/30 and a stereoacuity of at least 140" of arc as measured with the Diagnostic Series of the Keystone Telebinocular.

(4) Procedure. Each *O* made judgments of the size and distance of a rectangle at each of the four distances with each of the three interpupillary devices. Throughout this study, the instructions were such as to produce 'objective' judgments,[11] *i.e.* the *O*s were asked to reproduce physical extents either visually or kinesthetically. The order in which the different judgments occurred was systematically varied among the *O*s.

(a) Distance judgments. Before making the distance judgments, each *O* was permitted 10 practice throws down the throwing alley. *O* was instructed that the floor of the throwing alley was at the same height as that of the bottom of the rectangles, and the significance of this for his throwing was discussed. Following this, *O* looked through the interpupillary device and threw a dart as far down the throwing alley as the rectangle appeared to be from himself. The process of throwing darts was repeated for the same rectangle and the same *O* until five successful throws (throws in which the dart stuck into the floor of the throwing alley) had been completed. *E* measured the distances from the position of *O* to the five darts with the average of these distances constituting a throwing score. Another rectangle was then presented at a different distance and the process was repeated until throwing scores had been obtained at each of the four distances.

(b) Absolute size-judgments. The absolute size-judgments were made with the hand adjustable apparatus to apparently duplicate the 30.5 cm. width between the black rods presented in a frontoparallel plane at a distance of 152.4 cm. from *O*. Each *O* repeated this adjustment four times. Two of the adjustments were initiated with the hands close together, and two were begun with the hands far apart.

In making the relative size-judgments, *O* signaled when *E* had adjusted the lateral distance between the black rods in the visual alley to appear to be physically equal to the width of the particular rectangle. This was repeated three times for each *O* for each distance of the rectangle.

RESULTS

(1) Judgments of size and distance. The means of the judgments obtained are given in Table I. From the average absolute size-adjustments,

---

[11] Carlson, *op. cit.*, 199.

it is clear that frontal sizes in the visual alley did not appear equal for the different values of base. Using the *t*-test, the average values of perceived absolute frontal size from the increased base (*a*), the normal base (*b*), and the decreased base (*c*) were significantly different at the 2% level ($t_{c\text{-}b} = 2.90$, $t_{b\text{-}a} = 2.96$ and $t_{c\text{-}a} = 3.76$). Frontal sizes appeared largest for the decreased base and smallest for the increased base. From the perceived absolute size-scores at the one distance, and the perceived relative size-scores at the several distances, perceived absolute size-scores were computed for all the distances. This was accomplished by multiplying the relative size-scores by the appropriate ratio of the absolute size-scores. To convert the relative size-scores to perceived absolute frontal sizes (S') for the decreased, normal, and increased base, the relative size-scores were multiplied by 40/31.9, 31.9/31.9, and 28.5/31.9, respectively (see the right-hand portion of Table I). The throwing scores were taken directly to be measures of perceived absolute distance (D').

TABLE I

AVERAGE RESULTS (IN CM.) OF THE SIZE AND DISTANCE JUDGMENTS
($N = 12$)

Target distance (in cm.)	Throwing scores			Relative size-adjustments			Absolute size-adjustments		
	decreased base	normal base	increased base	decreased base	normal base	increased base	decreased base	normal base	increased base
152.4	180.8	117.6	100.9	7.8	8.0	8.2	40.0	31.9	28.5
304.8	289.3	201.8	186.4	17.6	18.6	22.3			
457.2	376.3	302.2	249.3	24.2	26.0	29.7			
609.6	451.1	401.6	338.7	28.6	30.3	34.0			

Fig. 2 shows the relation between the average throwing scores (D') and the perceived absolute size-scores (S') with S' computed in the manner described above. Since the visual angle (θ) of the rectangles was a constant in this study, S' and D' according to Equation [1] should be proportional. It is immediately clear from Fig. 2 that for each value of the base, S' is an increasing though not necessarily a linear function of D'. The curve labeled "Decreased Base" appears to be approximately linear but not to pass through zero, as Equation [1] would require. Neither of the curves labeled "Increased Base" nor "Normal Base" appear to fit Equation [1].

If Equation [1] is to apply successfully to the data of this study, S'/D' must be a constant. Fig. 3 shows how the average computed value of S'/D' varied as a function of physical distance for each value of base. The significance of the changes in the value of S'/D' was tested by an analysis of variance. The ratios of perceived absolute size to perceived absolute distance varied with the *O*s ($F = 6.38$, $df = 11$ and 66, $p < 0.01$), physical distance ($F = 17.70$, $df = 3$ and 33, $p < 0.01$), and base ($F = 4.68$,

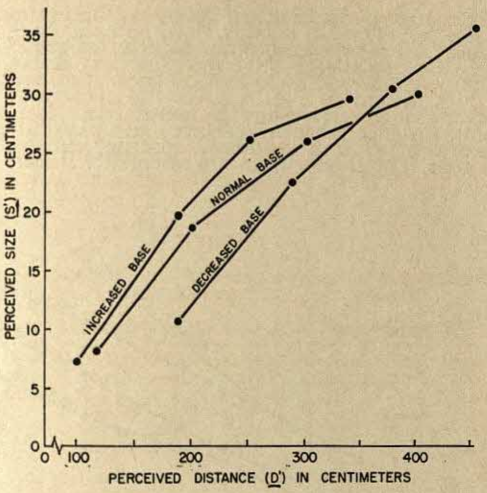


FIG. 2. RELATION BETWEEN PERCEIVED ABSOLUTE SIZE (S') AND PERCEIVED ABSOLUTE DISTANCE (D') FOR THE THREE BASES

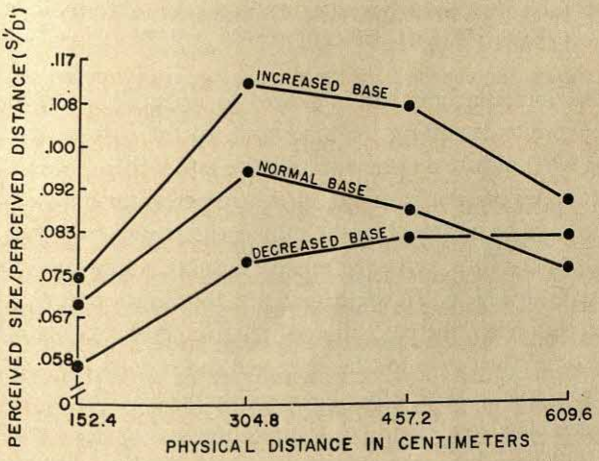


FIG. 3. VARIATION IN S'/D' AS A FUNCTION OF BASE AND DISTANCE

$df = 2$ and 22, $p < 0.05$) with all the first order interactions significant at the 5% level.

(2) Changes in perception as a function of base. Some of the perceptual consequences of changing the effective interpupillary distance in this study are shown in Figs. 4 and 5. Fig. 4 gives the average relation between perceived absolute distance (thrown distance) and physical distance for the three values of base. The three curves are essentially linear. It seems, from Fig. 4, that increasing the base appeared to make all the objects appear closer to *O* while decreasing the base made all the objects appear farther away. The effect of base on perceived relative rather than perceived absolute distance can also be seen from Fig. 4, where perceived relative depth may be considered to be measured by the difference between values of perceived absolute distance. Thus, perceived relative distance is measured by the slope of the curves in Fig. 4. To the extent that the three curves of Fig. 4 are parallel to each other, the perception of relative depth was not affected by a change in base. The curves for the decreased and normal base conditions are reasonably parallel to each other but possibly not parallel to the curve for the increased base condition. The dashed line in Fig. 4 shows the results which would have been obtained if perceived absolute distance had been equal to physical distance. The solid curves are not parallel to the dashed curve. Since it would be expected that the curves for all three values of base in Fig. 4 should pass through zero on both axes, it follows that for objects nearer than 152.4 cm. these curves can not be parallel to each other.

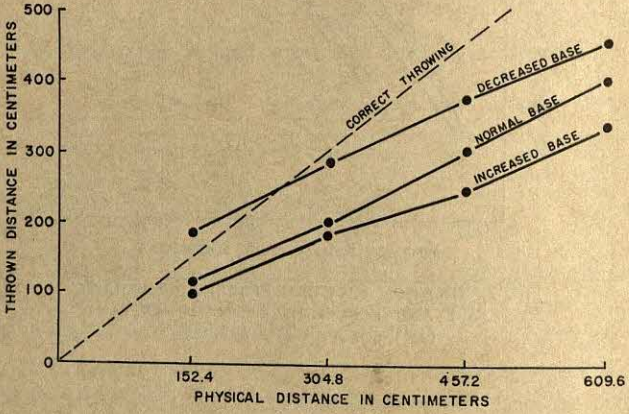


FIG. 4. PERCEIVED ABSOLUTE DISTANCE AS A FUNCTION OF PHYSICAL DISTANCE FOR THE THREE VALUES OF THE BASE

Fig. 5 indicates that perceived relative size seems to vary somewhat with physical distance and (perhaps somewhat more systematically) with base. The dashed curve in Fig. 5 gives the results which would have been obtained if the distance which *O* adjusted between the black rods had been equal to the physical width of the black rectangles, that is, the dashed curve represents perfect frontal size-constancy. A slope greater than this curve of perfect size-constancy would represent 'over' size-constancy while

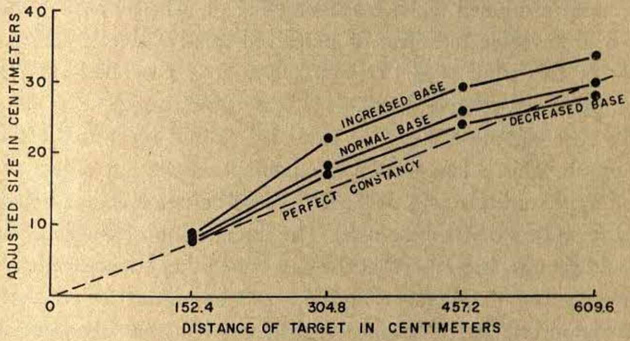


FIG. 5. PERCEIVED RELATIVE SIZE AS A FUNCTION OF PHYSICAL DISTANCE FOR THE THREE VALUES OF THE BASE

a lesser slope would represent 'under' size-constancy. From Fig. 5 there was a slight 'over' constancy between 152.4 and 304.8 cm., essentially 'perfect' constancy between 304.8 and 457.2 cm. and 'under' constancy between 457.2 and 609.6 cm.

DISCUSSION

(1) Perceptual effects of changing base. The most striking perceptual consequences of the different interpupillary distances occurred with respect to perceived absolute size and distance rather than perceived relative size and distance. In general, it was as though an increase in base made all the objects (between 152.4 and 609.6 cm.) appear to move closer by a constant amount and to decrease in absolute size, while a decrease in base made all the objects appear to move farther away by a constant amount and to increase in absolute size. How can these results be explained? It is clear that the convergence of the eyes was not an important factor since it remained constant for the average distance of the target in spite of changes in base. Also, these results would not be anticipated from the geometry of Fig. 1. From this figure, it would be expected that changing the base would result

in a change in perceived relative depth. Instead, within the distances used in this study, both perceived absolute size and perceived absolute distance were modified, while perceived relative depth showed little if any systematic change. The differences in the effects of base magnification or minification on perceived absolute and relative extents is a striking verification of the need to distinguish between perceived absolute and relative extents in both concept and mensuration.

A tentative explanation of the effect of base upon perceived absolute distance was mentioned when discussing Fig. 4. The slopes of the curves in Fig. 4 are average measures of perceived relative depth. Thus, for the distances of this study, *i.e.* 152.4 to 609.6 cm., base had no systematic effect on perceived relative depth. It is reasonable to expect, however, that the three base curves of Fig. 4, if extended, would meet at the zero point on the graph. This is based upon the assumption that an object very close to *O* (at approximately zero distance) would be perceived as being at that distance regardless of the base used. The intersection of the three curves at zero would require, however, that the curves be other than parallel for some distances less than 152.4 cm. from *O, i.e.* that base must differentially affect the perception of relative distance at distances less than those used in this experiment. It is also reasonable to expect that successive perceived relative depths would summate to produce a perceived absolute distance. It follows from this argument (and in agreement with the obtained results) that the perceived *absolute* distance of the rectangle, at 152.4 cm., (and all greater distances) would be different for the different bases.

The usual explanation of the effect of interpupillary distance on perceived frontal size would be in terms of perceived absolute distance, *i.e.* as the rectangles appeared to be at greater distances for lesser values of base, they should have appeared to be of a larger absolute size. As indicated by Fig. 2, however, such an explanation is not entirely satisfactory. A different and perhaps more appropriate explanation can be given in terms of the perceptual interrelation of frontal and depth extents.

It has been demonstrated that, under conditions in which binocular disparity is the only cue to perceived relative depth, a change in perceived frontal size (per unit of angular frontal size) can induce a change in the perceived depth associated with a constant binocular disparity.[12] The converse might also occur, as seemed to be the case in the present study. In the present study, cues in addition to binocular disparity were present to determine perceived relative depth. Perhaps as a consequence of this, a change in binocular disparity, as a result of a change in base, might be interpreted by *O* as a change in perceived frontoparallel size rather than as a change in perceived relative depth. This would mean that the failure of perceived relative depth to increase proportionally with an increase in binocular disparity was responsible for the change in the perceived size of frontoparallel extents found in the present study. From this point of view, the lack of change in perceived relative depth with different bases would be expected to result in perceived frontoparallel size increasing for a decrease in interpupillary base and decreasing for an increase in interpupillary base. Since, however, a complex of cues occurred in the present study, it is not expected that this point of view can predict the magnitude as well as the direction of the changes in perceived frontoparallel size shown in Table I.

---

[12] Gogel, *op. cit.,* 1959, 1-16, *op. cit.,* 1960, 119-131; 179-192; 257-269.

It may be noted that the discussion of the relation between cues of binocular disparity and the perceived size of frontally presented objects is also an expression of a relation between perceived size and distance. It is not, however, the hypothesis of the invariance of the ratio of perceived size to perceived distance of Equation [1] since it involves relative not absolute perceived distance.

(2) Hypothesis of invariance. In testing the hypothesis of invariance in the form of Equation [1], it was essential that the *measurements* of perceived absolute size and perceived absolute distance were actually proportional to perceived absolute size and perceived absolute distance. If this were not so, this study would not constitute a valid test of Equation [1]. In other words, the hypothesis of invariance has been validly tested *only* if

$$S' = K_1 S'_{\mathrm{m}} \text{ and} \qquad [4]$$

$$D' = K_2 D'_{\mathrm{m}}, \qquad [5]$$

where S' and D' are perceived absolute size and perceived absolute distance, K_1 and K_2 are constants, and S'_m and D'_m are the measures of perceived absolute size and perceived absolute distance obtained from the size-adjustments and throwing scores. Evidence supporting the position that the conditions of Equations [4] and [5] were satisfied in this experiment is available.

There are strong indications that the magnitude of the kinesthetic (absolute) size-adjustments were proportional to perceived absolute size After finishing the main portion of this study, 12 *O*s (9 of whom had been used in the main portion of the study) were used to conduct a direct test of the relation between the method of kinesthetic size-adjustment and the perception of absolute size. For this purpose, several lateral separations (11.4, 22.9, 34.3, and 45.7 cm.) between the black rods in the visual alley were presented in random order. Using the hand adjustable apparatus, each *O* kinesthetically adjusted the lateral distance between the aluminum rods at the *O*'s position until it appeared kinesthetically to be physically equal to the lateral distance between the black rods. As before, the aluminum rods of the hand adjustable apparatus were not visible during the kinesthetic adjustments. The kinesthetic adjustment was repeated four times for each of the four lateral separations of the black rods. Since the black rods were at a constant distance from *O* (152.4 cm.), physical and perceived extents between the rods would be expected to be proportional. If, therefore, the magnitudes of the hand adjustments required to apparently duplicate the lateral distances between the black rods were proportional to the physical distances between the black rods, it can be concluded that the hand adjustments were proportional to perceived size. It is evident from Fig. 6 that this was the case. Thus, the kinesthetic size-adjustments (the hand adjustments) satisfy Equation [4].

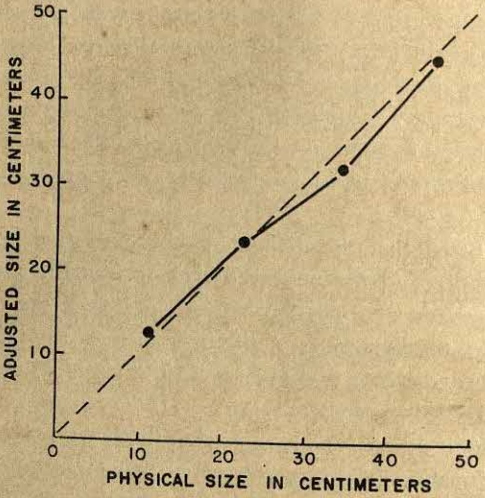


FIG. 6. RELATION BETWEEN THE KINESTHETIC SIZE AND PHYSICAL SIZE AT A CONSTANT PHYSICAL DISTANCE

Perceived absolute sizes for the rectangles at the 152.4, 304.8, 457.2, and 609.6 cm. distances for the different bases were obtained by multiplying the perceived relative size-adjustments by a ratio of the perceived absolute size-adjustments made with the different bases at the 152.4 cm. distance. It is necessary, therefore, to examine whether the relative size-adjustments were proportional to perceived relative size. There are two possible reasons why measurements of perceived size and the actual perceived size may differ. First, if the measurement involves a motor adjustment by *O*, some motor error may occur and may vary as a function of the magnitude of the adjustment. Secondly, the modality in which the measurement is made might not be proportional to perceived visual size. Apparently, from Fig. 6, neither of these kinds of errors occurred with the kinesthetic size-adjustments and it is evident that neither kind could have occurred with the visual size-adjustments. With the visual (relative) size-adjustments, motor errors did not occur since *E* made the actual adjustments in response to the commands of *O*. The second kind of error was not present since visual size cannot have a disproportionate relation to itself. On logical grounds, therefore, the visual (relative) size-adjustments were proportional to perceived relative size.

Empirical evidence is also available indicating that the relative size-adjustments were proportional to perceived relative size. In a previous study, the kinesthetic and visual methods of measuring perceived size were compared under conditions showing practically perfect frontal size-constancy.[13] It was found (Fig. 5 of the previous study) that the two methods gave proportional results. Thus, if the kinesthetic size-adjustments satisfy Equation [4], as the present study suggests, the relative size-adjustments give scores which are proportional to perceived relative size.

There is evidence that Equation [5] is also satisfied in the present study. The data from the 'normal base' in Fig. 4 lie approximately on a straight line passing through zero. This is evidence that the throwing scores were proportional to physical distance; but the situation using the normal base approaches what might be called a 'full-cue' situation; *i.e.* a situation in which a variety of cue-systems are available to determine perceived extents such that veridical depth perceptions are most likely to occur. It may be assumed that, with the normal base condition, perceived and physical depth were proportional. It follows that the throwing scores obtained in the present study were directly proportional to perceived absolute distance as required by Equation [5]. In previous studies in which similar full-cue situations were investigated, throwing scores were proportional to physical distance in one case, and were linear but probably not proportional in the other.[14]

The evidence, therefore, is that, in the present study, Equations [4] and [5] are satisfied and that, consequently, Equation [1] has been validly tested.

---

[13] Gogel, *op. cit.*, 1962, 91-104.

[14] Gogel, Hartman, and Harker, *op. cit.*, 1-16; Gogel, *op. cit.*, 1960, 119-131.

It is evident from Figs. 2 and 3 that the results from the present study do not support Equation [1]; but, tests of the hypothesis of invariance less stringent than Equation [1] can also be made. One such test is that S' is a linear but not necessarily a proportional function of D'. In order for this hypothesis to be validated, all three curves of Fig. 2 should be linear and coincident, though it is not required that they pass through zero. Only one of the curves is linear. The fact that the curves are not coincident tends to reject even the hypothesis that S' has some fixed relation to D' for a given *O* for different viewing conditions.

In testing Equation [1], θ need not be constant. If θ is varied, S'/D' according to Equation [1] should vary proportionately. Thus, Equation [1] assumes that the retinal size θ can have consequences for the perception of absolute size (or distance). The term θ does not involve a ratio of retinal sizes. It is an absolute retinal size. As discussed in a previous study, there is evidence, however, for doubting the assumption that the magnitude of an absolute retinal size has any perceptual consequences.[15] This evidence, therefore, is directly opposed to the hypothesis that the ratio of perceived size and perceived distance is invariant. Equation [1] would also predict that perfect depth-constancy occurs whenever perfect frontal size-constancy is present. It has been found, however, that under certain conditions, when perfect frontal size-constancy was present perfect depth-constancy, using the binocular disparity cue, clearly did not occur.[16] This evidence, therefore, is also in opposition to the hypothesis of invariance.

The value of K in Equation [1] would be expected to be different for different *O*s. Rump indicates that this constant is sometimes regarded as a result of the situation and of the observer.[17] If K varies as a function of situations (or within a given situation) Equation [1] loses much of its generality. Kuroda suggests that the relation between S' and D' is a power function with the value of the exponent determined by the amount of size-constancy present.[18] This raises the question of whether the introduction of additional meanings or additional terms into Equation [1] is necessary in order to make the relation valid. In other words, can an equation be written in which S' can be predicted from D' in all situations without requiring any change in the constants of the equation? From the present study the answer to this question is 'no.' Any equation relating S' and D', if it is to apply to all viewing conditions, must contain terms (other than those of θ, S', and D') which vary as a function of the viewing conditions.

---

[15] Gogel, *op. cit.*, 1962, 91-104.

[16] Gogel, *op. cit.*, 1960, 257-269.

[17] Rump, *op. cit.*, 111.

[18] T. Kuroda, Experimental studies on size constancy, *Kyushu Psychol. Stud.*, 2, 1961, 59-102.

SUMMARY

The possible interrelation of perceived absolute size and perceived absolute distance was studied using three values of effective interpupillary distance (base). For this purpose, independent measures of perceived size and perceived distance were obtained for objects subtending a constant-visual angle at 152.4, 304.8, 457.2, and 609.6 cm. The devices used to provide the different interpupillary distances were such that the convergence of the eyes to a distance of 381.0 cm. was constant for all values of base. The results of the study indicate that perceived absolute size and perceived absolute distance are positively correlated. In disagreement, however, with the hypothesis of the invariance of the ratio of perceived size and perceived distance, the ratio of perceived absolute size to perceived absolute distance varied as a function of both base and physical distance.

The different interpupillary devices used in this study resulted only in changes in the binocular disparity cues to relative depth. Perceptually, however, for the distances used, a change in perceived absolute size and distance occurred with perceived relative size and distance showing less effects. Possible explanations for these results were considered.

THE SEMANTIC DIFFERENTIAL AND THE STRUCTURE OF INFERENCES AMONG ATTRIBUTES

By DANIEL KAHNEMAN, The Hebrew University, Jerusalem

Ever since the introduction of the semantic differential as a tool for the measurement of connotative meaning, interest has centered on the correlations among the adjectival scales of which the instrument consists. One of the main findings reported by Osgood, Suci, and Tannenbaum was the discovery of the three factors, evaluation, potency and activity, which together account for most of the covariance of scales.[1] These factors have appeared quite regularly in a long series of investigations, which varied the concepts rated, the scales themselves, and even the language of the respondents.[2] A subsidiary focus of research has been the investigation of specific concepts, or groups of concepts, for which the usual correlational structure breaks down. These cases have usually been interpreted by the assumption of additional cognitive processes, by which the meaning of scales is changed in correspondence to the meaning of the concepts rated. While elaborate statistical techniques have often been used in this research, there has been relatively little stress on the development of a precise algebraic model to aid in the evaluation of obtained correlational results. The present paper reports an attempt to develop such a model and elaborate some of its consequences.

A simple model can be formulated for semantic ratings, patterned after existing models of rating-scales.[3] The rating s_{ijk} of concept J by individual i on scale k is a sum of three components:

$$s_{ijk} = T_{jk} + C_{ik} + d_{ijk} \qquad [1]$$

T_{jk} is the 'true score' of concept j on scale k. It is obtained by averaging the ratings given to concept j by a large number of judges. The semantic norms com-

---

* Received for publication April 21, 1962. This paper is abstracted from parts of a dissertation submitted to the Graduate Division of the University of California, Berkeley, in partial fulfillment of the requirements for the Ph.D. degree under the title *An Analytical Model for the Semantic Differential* (1961). The writer is greatly indebted to Drs. E. E. Ghiselli, Susan M. Ervin, and J. Block for their generous help and advice. Most of the analyses were performed by the Computer Center of the University of California.

[1] C. E. Osgood and G. J. Suci, Factor analysis of meaning, *J. exp. Psychol.*, 50, 1955, 325-338. C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, *The Measurement of Meaning*, 1957, 342.

[2] Osgood, The cross-cultural generality of visual-verbal synesthetic tendencies, *Behav. Science*, 5, 1960, 146-169; Studies on the generality of affective meaning systems, *Amer. Psychol.*, 17, 1962, 10-28.

[3] J. P. Guilford, *Psychometric Methods, 2nd ed.*, 1954, 281.

piled by Jenkins, Russell, and Suci are a collection of such true scores.[4] C_{ik} is the constant deviation of Observer *i* on scale *k*. It is obtained by averaging *O*'s deviations from true scores over a large number of concepts. The remainder, d_{ijk}, is a specific deviation on a particular rating. By obtaining repeated ratings of concepts by the same *O*s, it is possible further to subdivide this component into a consistent idiosyncrasy of an *O* in rating a particular concept and into a momentary 'error.'

This model has immediate implications for the correlational analysis of semantic ratings. Several coefficients of correlation among scales can be defined, corresponding to the components of the judge's rating. These are:

(1) The correlation of true scores, r_T, which is defined within a universe of concepts, *e.g.* concepts which are generally rated LARGE tend to be rated STRONG. The structure of these correlations is obviously manipulable by the choice of concepts. It is quite possible, for instance, that the true scores of political figures are such that GOOD politicians are also STRONG and ACTIVE.[5] The same may not be true if the concepts to be rated are lyrical poems. Carroll has argued that r_T is the most appropriate method of computing correlations among semantic scales.[6]

(2) The correlation of constant deviations, r_C, which is defined for a population of judges, *e.g.* raters who generally exaggerate the strength of objects are likely to exaggerate their size as well. This kind of correlation has not been discussed in the literature of the semantic differential.

(3) The average correlation of deviations, $\bar{r}_d$. This coefficient may be defined in two ways: as an average over raters of r_{di}, the correlations of deviations for a particular rater *i* over a set of concepts; as an average over concepts of r_{dj}, the correlation of deviations in the ratings of each concept *j* by a group of judges. The two methods of averaging coefficients (after transformation of each coefficient to Fisher's *z*) always lead to nearly identical results. Regardless of which method is used, $\bar{r}_d$ indicates a consistent tendency on the part of *O*s to produce correlated errors of judgment. Thus, if an *O* deviates from consensus and his own constant error in rating a concept as STRONGER than others do, he is likely to deviate in judging it to be LARGER than it is. Such correlated errors may be produced if both ratings are partly determined by the operation of a single fallible mechanism. This view is quite consistent with Osgood's interpretation of semantic factors as mediational mechanisms of response. Alternatively, we may interpret correlated errors as due to a system of inferences: if the object rated is STRONG, it is unlikely to be very DELICATE or very SMALL. If the rater corrects a set of preliminary estimates according to such inferences, his errors on the various dimensions will be correlated.

The three main components of semantic ratings are independent by definition.[7]

---

[4] J. J. Jenkins, W. A. Russell, and G. J. Suci, An atlas of semantic profiles for 360 words, this JOURNAL, 71, 1958, 688-699.

[5] Osgood, Suci and Tannenbaum, *op. cit.*, 114-115.

[6] J. B. Carroll, Critique of the semantic differential, *Amer. Psychologist*, 14, 1959, 382.

[7] The three components are necessarily independent over the entire set of ratings. The more demanding assumption is made here that the specific deviations are independent of the true scores of concepts for each individual separately, and of the constant deviations of raters for each concept separately. A modification of the first of these assumptions will be discussed below. On the general problem of the correlation between two variables in two orthogonal planes of a cube of data, see George Mandler, Stimulus variables and subject variables: A caution, *Psychol. Rev.*, 66, 1959, 145-149.

Consequently, the various correlations defined are also mathematically independent and may provide answers to distinct questions of theoretical interest. By contrast, the index of correlation chosen by Osgood and his co-workers cumulates variance and covariance from the three basic components. This coefficient will be referred to as the cumulative correlation. It is given by:[8]

$$r_{\text{cum}} = \frac{I\sigma_{Tk}\sigma_{Tl}r_T + J\sigma_{Ck}\sigma_{Cl}r_C + \sum_i \sigma_{dik}\sigma_{dil}r_{di}}{\sqrt{I\sigma_{Tk}^2 + J\sigma_{Ck}^2 + \sum_i \sigma_{dik}^2}\sqrt{I\sigma_{Tl}^2 + J\sigma_{Cl}^2 + \sum_i \sigma_{dil}^2}}$$

$$\mathrm{R} = \frac{100[\sigma^2(1+2) - (\sigma_1^2 + \sigma_2^2)]}{2\sigma_1^2},$$

It is apparent from Formula [2] that the value of the cumulative correlation depends on the values of the primary coefficients defined above, weighted by the respective variances of the components. The cumulative correlation is also affected by the relative number of raters and concepts in a particular investigation. This coefficient is interpretable without ambiguity only when the correlations within the three orthogonal components are closely similar. As will be shown below, this condition is probably fulfilled in most cases.

The two studies described in this paper attempted to answer several questions raised by the present conception of the semantic differential. Of main interest in this investigation were the characteristics of correlations among deviations, or errors, in semantic judgments, and the identification of these correlations as representing a system of inferences among attributes.

METHOD

Sixty *Os*, students in psychology at the University of California at Berkeley, rated 40 concepts on a semantic differential consisting of 12 scales. The concepts were chosen from the norms published by Jenkins *et al.*,[9] who reported the mean ratings (true scores) on 20 scales for 360 concepts, each of which had been judged by 30 students at the University of Minnesota. Four sets of 10 concepts each were so selected as to produce different correlations of true scores among some central scales. Set A was so selected as to induce a negative correlation between the scales WEAK-STRONG and SLOW-FAST. The following concepts were included in Set A: *Peace, Lake, Dawn, Birth, Curved, Fear, Flea, Sin, Criminal, Thief*. Set B was so composed as to induce a highly positive correlation of true scores between STRONG and FAST. It included the following concepts: *Sword, War, Sex, Trouble, Sister, Lady, Grief, Gloomy, Dim, Beggar*. Set C included concepts among which the correlation of true scores between STRONG and GOOD was highly positive: *Courage, Health, Money, Debate, Obvious, Bottom, Obscure, Broken, Deformed, Inferior.* Finally, Set D included concepts among which the true scores of STRONG and GOOD were negatively correlated: *Danger, Hit, Capital Punishment, Liquor, Hurried, Somber, Pliable, Girl, Puppies, Jelly.*

---

[8] The symbol σ_{Tk} designates the standard deviation of the true scores on scale *k*. Similarly, σ_{Ck} and σ_{dik} stand for the standard deviation of constant deviations (over individuals) and for the standard deviation of specific deviations (over concepts) for individual *i* on scale *k*, respectively. *I* is the number of *Os*, *J* is the number of concepts. An algebraically equivalent result is obtained by computing and cumulating r_{dj} and σ_{dj} instead of r_{di} and σ_{di}.

[9] Jenkins, Russell, and Suci, An atlas of semantic profiles for 360 words, this JOURNAL, 71, 1958, 688-699.

The following scales were selected as representative of the major factors described by Osgood *et al.*: CRUEL-KIND, ACTIVE-PASSIVE, SOFT-HARD, STRONG-WEAK, COLORLESS-COLORFUL, ANGULAR-ROUNDED, SAD-HAPPY, FAST-SLOW, BAD-GOOD, MASCULINE-FEMININE, CALM-EXCITABLE, SMALL-LARGE.

The instructions to the *O*s were adapted from the instructions used by Osgood and his associates.[10] At intervals of 5-8 days after the first session, 20 of the *O*s appeared for a second session in which one of the tasks was to rate again the concepts included in Set A. The instructions for the second session mentioned the fact that some of the concepts would appear familiar, and emphasized the requirement that the ratings represent first-feeling reactions and be made independent of each other and of previous ratings.

Thirty *O*s participated in the second study, in which ratings of inferences among attributes were obtained. All 30 *O*s completed the ratings of the 10 concepts of Set A before turning to the second procedure. Twenty of the *O*s had already participated in the first study, approximately one week previously.

Ratings of inferences were obtained as follows: the *O*s were given a booklet of 24 pages, laid out according to the standard format of the semantic differential. The 24 adjectives defining the 12 dimensions used in the first study now appeared successively as concepts. The *O*s task was defined as a 'game' in which, given only one property of an unknown object, indicated at the top of the page, they were to estimate the likelihood that the object would have one or the other of the traits defining the 12 scales on the page. Inferences were obtained from each of the 24 adjectives onto the 12 scales used in this study.[11]

RESULTS

Partition of the variance of semantic ratings. Table I shows that the variance of the true scores of the concepts accounts for an appreciable part of the variance of ratings. Because of the deliberate choice of widely varying concepts in this experiment, the variance of true scores is larger than might generally be expected. In several sets of data previously analyzed, the variance of true scores was only about half as large as the mean variance of deviations. The constant errors of raters are consistently small, implying agreement on the average in the use of scales.[12]

---

[10] Osgood, Suci, and Tannenbaum, *op cit.*, 82-84.

[11] The instructions for the second procedure are presented in the original report, 38-39.

[12] Similar results have been reported by C. F. Gross, Intrajudge consistency in ratings of heterogeneous persons, *J. abnorm. soc. Psychol.*, 62, 1961, 605-610. This investigator found that only about 3% of the total variance of interpersonal ratings was due to differences among raters in their constant deviations.

From the results reported in the last column of Table I, it appears that little more than a third of the variability among judges in the rating of particular concepts is reliable over one week.[13] Thus, most of the variance of deviations consists of relatively short-lived fluctuations around the true error of measurement.

Correlations among deviations. The correlations among scales were computed by the four methods indicated above in the four sets of concepts. The matrix $\bar{r}_d$ was derived by averaging corresponding coefficients d_{di} computed separately for each of the 60 judges, after transformation to Fisher's z. The average is justified by the homogeneity of coefficients among individuals: the variance of corresponding coefficients over individuals is only slightly larger than would be expected by chance, according to the theoretical distribution of Fisher's z. As was indicated above, the matrix $\bar{r}_d$ obtained in this manner is almost identical to the matrix which would be obtained by averaging coefficients r_{dj} computed separately for each concept.

TABLE I

VARIANCES OF TRUE SCORES (σ_T^2), OF CONSTANT DEVIATIONS (σ_C^2), AND MEAN VARIANCE OF DEVIATIONS ($\bar{\sigma}_d^2$). (PERCENTAGE OF ORIGINAL VARIANCE OF DEVIATIONS REPEATABLE AFTER ONE WEEK, R)

Scale	σ_T^2	σ_C^2	$\bar{\sigma}_d^2$	R^*
KIND-CRUEL	1.74	0.07	1.02	35
ACTIVE-PASSIVE	1.78	0.17	1.74	18
SOFT-HARD	1.98	0.08	1.40	36
STRONG-WEAK	1.51	0.18	1.82	45
COLORLESS-COLORFUL	1.06	0.19	1.76	32
ANGULAR-ROUNDED	1.31	0.12	1.60	36
SAD-HAPPY	2.26	0.10	1.12	25
FAST-SLOW	1.31	0.11	1.59	33
BAD-GOOD	2.49	0.08	1.11	41
MASCULINE-FEMININE	1.52	0.08	0.96	36
CALM-EXCITABLE	1.55	0.14	1.64	52
SMALL-LARGE	0.93	0.18	1.35	36

* The values in this column were computed according to the following formula:

$$R = \frac{100[\sigma^2(1+2) - (\sigma_1^2 + {}_2^2)]}{2\sigma_1^2},$$

where σ_1^2 is the mean variance of deviations on the first test, σ_2^2 is the mean variance of deviations on the second test and $\sigma^2_{(1+2)}$ is the mean variance of the sums of corresponding deviations for the two occasions.

Table II shows that there are considerable differences in the average absolute values of the different coefficients. As Carroll has already noted.[14]

---

[13] These results are consistent with the findings reported by W. T. Norman, Stability-characteristics of the semantic differential, this JOURNAL, 72, 1959, 581-584.

[14] Carroll, *op. cit.*, 382.

the coefficients of correlation among true scores are the highest. The correlations r_C among constant deviations are lower, and the values for $\bar{r}_d$ are lower still. The cumulative correlation, being a composite, takes intermediate values. All differences are highly significant within each set of concepts.

The most significant aspect of Table II is the fact that substantial correlations $\bar{r}_d$ are found among specific deviations. Specific deviations from the true scores include both consistent idiosyncrasies in the rating of a particular concept and momentary deviations from the true score. It may appear likely that the observed correlations $\bar{r}_d$ are entirely due to consistencies in the more lasting deviations, which are merely attenuated by random uncorrelated error. On this assumption, averaging the ratings given by each *O* on two occasions should produce a marked increment in the size of r_d, since the reliability of single ratings is low. This is not the case. For 20 *O*s rating the 10 concepts of Set A, the mean absolute value of the coefficients $\bar{r}_d$ was 0.166 on the first test, 0.152 on the second, and 0.179 for the averaged ratings. Thus, even relatively momentary deviations from true scores are interrelated to a substantial degree.

TABLE II

MEAN ABSOLUTE LEVEL OF THE CORRELATIONS AMONG SCALES IN FOUR SETS OF CONCEPTS

Correlation	Concepts			
	A	B	C	D
r_T	.600	.564	.622	.579
r_C	.218	.190	.227	.211
$\bar{r}_d$	.174	.178	.214	.211
r_{cum}	.330	.372	.346	.362

This conclusion is at variance with an analysis by Osgood, Suci, and Tannenbaum, who concluded that the error of measurement in the semantic differential is uncorrelated.[15] Osgood *et al.* computed *differences* between the ratings given to a particlular concept on each scale for successive occasions of testing. They correlated test-retest differences between scales representing the same factors. Very low correlations were obtained for intervals of 3-30 min. more substantial correlations appearing for intervals of 1 day to 3 weeks. Algebraic analysis shows that these results imply that the correlation between the error of measurement on Scale k at time t_0 and the error of measurement on Scale l at time t_1 is a decreasing function of the interval $(t_1 - t_0)$. For this to be the case, the correlation between errors at time t_0 must have been higher than zero. As a matter of fact, the correlation between errors at a particular time must have been at least as high as the correlations between changes for the longest intervals in Osgood's study. The main conclusion of this section, that even momentary deviations from true scores are intercorrelated, is thus actually supported by the data presented by Osgood *et al.*

---

[15] Osgood, Suci, and Tannenbaum, *op. cit.*, 135-137.

The correlational structures in the cognitive system and in ecology. Since the three components in the variance of the semantic differential are mathematically independent, it is possible for the deviations on the various scales to be correlated in an entirely different way than the true scores of concepts on these scales. This is unlikely, however, if the deviations represent inferences and the true scores represent the ecology from which such inferences must have been drawn. The semantic differential is well suited to a study of this question, since the same data provide estimates of the correlations among attributes both in the environment (r_T) and in the cognitive system $(\bar{r}_d)$.

TABLE III

CORRELATIONS AMONG CORRELATIONAL MATRICES IN FOUR SETS OF CONCEPTS

Correlation	Concepts			
	A	B	C	D
$\bar{r}_d - r_T$	.73	.89	.91	.89
$r_T - r_C$	.79	.85	.88	.68
$\bar{r}_d - r_C$	.57	.78	.86	.41

To assess the similarity between the correlational structure of true scores and of deviations, intercorrelations were computed between corresponding entries in the various correlational matrices. This index reflects the similarity in the *pattern* of relationships, but remains unaffected by the absolute magnitude of the coefficients. Table III includes the correlations between three of the correlational matrices obtained for each of the four sets of data: the correlation of true scores r_T; the correlation of constast deviations r_C; and the average correlation of deviations $\bar{r}_d$.

The similarity between the various correlational structures is certainly underestimated in Table III, since the sets of concepts were deliberately chosen in an attempt to create unusual relationships among true scores. The high level of the correlations under these circumstances lends support to the position that the cognitive system of the individual mirrors the relationships in the environment. As was indicated above, these data also provide justification for the cumulative coefficient used by Osgood.

A reinterpretation of interactions between concepts and scales. It is possible to compute the correlation between two scales within the ratings given by a group of *O*s to a single concept. Each such correlation is a composite of r_σ and r_{aj} for the particular concept *j*. The latter coefficient is weighted much more heavily in the composite (see Table I). Osgood has used this composite index in some of his studies of interactions among concepts and scales.[16] He reports different correlations for different concepts: in the ratings of *Adlai Stevenson,* for instance, PLEASURABLE is correlated with MASCULINE; but PLEASUREABLE goes with FEMININE for *My Mother*. Thus, *O*s who exaggerate Stevensons masculinity also tend to exaggerate his pleasantness, the converse being true in the rating of *My Mother*. Osgood's interpretation of such findings is that "in the process of human judgment, all scales tend to shift in meaning toward parallelism with the dominant (characteristic) attribute of the concept being judged."[17]

While the effects described by Osgood and his collaborators are undoubtedly present, it should be pointed out that they are of relatively limited magnitude. Complete matrices of intercorrelations among scales were obtained for each concept in the present study. Forty coefficients of correlation were therefore available for each pair of scales, each derived from one of the 40 concepts rated. The variance of corresponding coefficients was computed, within each set of 10 concepts, after transformation of each coefficient to Fisher's *z*. This variance should be large if interaction of concepts and scales is of importance in the semantic differential. Observed variances are small; however, the median value for the 66 pairings of scales is 0.028 for the 10 concepts of set A, for instance, as compared to 0.017 which is expected by chance on the basis of the theoretical distribution of Fisher's *z*.

Such instances of interaction as occur do not necessitate the postulation of specific changes in the meaning of scales when these are applied to particular concepts. A more parsimonious explanation is obtained by assuming that the specific deviation d_{ijk} is not entirely independent of the true score T_{jk}. Thus:

$$d_{ijk} = a_{ik}T_{jk} + e_{ijk}, \qquad [3]$$

a_{ik} may be interpreted as a 'constant of exaggeration' for judge i on scale k. An individual for whom $a_{ik} > 1$ will tend to make judgments more extreme than T_{jk}, but in the same direction with respect to the zero-point of the scale. An individual for whom $a_{ik} < 1$ will tend to place his judgments between T_{jk} and the zero-point.

The covariance of the deviations from consensus for a particular concept j is of the form:[18]

---

[16] Osgood, Suci, and Tannenbaum, *op. cit.,* 176-188. Osgood, *op. cit.,* 1962, 25.
[17] Osgood, Suci and Tannenbaum, *op. cit.,* 187.

$$\sum_i d_{ijk}d_{ijl} = \sum_i e_{ijk}e_{ijl} + \sum_i (a_{ik}T_{jk})(a_{il}T_{jl})$$

$$= \sum_i e_{ijk}e_{ijl} + T_{jk}T_{jl} \sum_i a_{ik}a_{il} \qquad [4]$$

since T_{jk} and T_{jl} are constants in the summation.

If we make the reasonable assumption that the constants of exaggeration on different scales are positively correlated, it is seen that the value of the covariance, and even its sign depend on whether T_{jk} and T_{jl} are on corresponding sides of their respective neutral points. If, for instance, the true scores of *Adlai Stevenson* indicate that he is both PLEASURABLE and MASCULINE, the correlation between the two scales will be positive. For the concept *My Mother,* which is PLEASURABLE and FEMININE, the sign of the correlation will be reversed. The other examples presented by Osgood *et al.* are equally consistent with this interpretation.[19] *My Mother* is certainly rated both SOFT and IMPORTANT, while *Engine* is rated HARD and IMPORTANT. The considerations presented above lead us to expect different signs for the correlation between these two scales, which is one of the striking examples cited by Osgood.

To test the validity of this interpretation, two pairs of scales were selected, for which the over-all correlation was approximately zero, and the variance of r_{dj} over concepts was large. The pairs selected were KIND (CRUEL)-WEAK (STRONG) and GOOD (BAD)-LARGE (SMALL). For each of the 40 concepts it was noted whether the true scores on the two scales were on corresponding sides of the neutral point. For both pairs of scales, positive correlations were most frequently found when T_{jk} and T_{jl} had similar signs (respectively, $X^2 = 9.55$, $p<0.005$; $X^2 = 13.19$, $p<0.001$). The interpretation of interactive effects in terms of individual differences in the expansiveness of raters is thus strongly confirmed.

The present analysis therefore suggests alternative interpretations for the two kinds of evidence which Osgood has mentioned for the prevalence of interactions in the semantic differential. The different correlations found for different *sets* of concepts, such as colors and words,[20] can be attributed to differences in the organization of true scores for these classes of concepts. Such results belong to an investigation of the ecology; no inferences can be drawn from them to changes in the meaning of scales, which we certainly a property of *O*'s cognitive system. The different correlations found for *single* concepts are due to individual differences among raters. The correlations which serve as evidence here are computed over *O*s, and any inference from such correlations to processes occurring within the individual is again necessarily doubtful.

---

[18] The subsequent development ignores the effect of constant deviations, for the sake of simplicity of exposition. This component has no significant differential effects in the comparison of correlations computed for different concepts, since it contributes fixed quantities to the numerator and to the denominator of all such correlations.

[19] Osgood, Suci, and Tannenbaum, *op. cit.,* 178.

[20] Osgood, *op. cit.,* 1962, 24.

Correlations among scales in different sets of concepts. The problem of choice of concepts in correlational studies has already been stressed in a critique of the semantic differential,[21] which emphasized the restricted number of concepts on which very general conclusions have often been based. The problem appears even more acute in the present framework, where the correlations of true scores r_T are said to be entirely predetermined by *E*'s selection of concepts to be rated. It is therefore of importance to establish the invariance of the correlations of deviations $\bar{r}_d$, which are theoretically independent of the concepts rated. Testing the stability of $\bar{r}_d$, under conditions where r_T varies is also a legitimate approach to the study of interactions among concepts and scales.

TABLE IV

MEDIAN COEFFICIENTS OF CORRELATION AMONG CORRELATIONAL MATRICES

Three arbitrary groups of 20 *S*s rated 4 sets of concepts. Correlational matrices were obtained by the different methods in each of the 12 sets of data. The 12 matrices obtained by each method were then intercorrelated. The table presents the median intercorrelations in several groups of comparisons.

Subjects	Concepts	Number of comparisons	r_T	$\bar{r}_d$	r_C	r_{cum}
Different	Same	12	.98	.86	.37	.98
Same	Different	18	.52	.80	.38	.65
Different	Different	36	.50	.79	.37	.64

The correlations among true scores were accordingly manipulated in the four sets of concepts, used in this experiment, in order to observe the concomitant changes in the correlations of deviations.

The sample of 60 raters was arbitrarily subdivided into 3 groups, providing for 12 sets of data (all raters had judged the 4 sets of concepts) which were analyzed separately. The correlational matrices derived by each method were compared across these 12 sets. It is clear from Table IV that the correlations of true scores varied considerably in the different sets of concepts, while the correlations of deviations were scarcely affected. The median intermatrix correlation for the $\bar{r}_d$ method decreases only from 0.86 when the concepts are the same to 0.80 when the concepts are different. This change is statistically significant, but is no larger than would be expected due to the 'constant of exaggeration' which entails a slight dependence of the deviations on the true scores. Under the same conditions, the median inter-matrix correlation for r_T decreases from 0.98 to 0.51. When concepts are the same, on the other hand, the correlations of true scores are more stable than the correlations of deviations. The correlations of constant deviations r_C are generally quite unstable: the constant errors are evidently too small to be reliably estimated from the rating of 10 concepts. The cumulative coefficient, which has been most widely used in research on the semantic differential, tends to be almost as stable as its most stable component.

---

[21] Carroll, Review of Osgood, Suci, and Tannenbaum's *The Measurement of Meaning, Language,* 35, 1959, 58-77.

Correspondence of inferences and correlations of deviations. The ratings of inference produced by the *O*s of the second study were used to derive a matrix of intercorrelations among scales. The *O*s had provided four sets of ratings of the relationship between each two scales, *e.g.* the inferences from KIND and CRUEL on the scale GOOD-BAD and the inferences from GOOD and BAD on the scale KIND-CRUEL. The ratio of each inference to the maximal inference possible, *i.e.* from a scale onto itself, is an estimate of a standardized coefficient of regression which is directly related to the coefficient of correlation. In a linear and symmetrical relationship, all four estimates should be equal. Although the conditions of linearity and symmetry frequently fail to be fulfilled,[22] an average of the four ratios provides an adequate measure of the relationship.

The entries in the matrix of correlations obtained in this manner were correlated to corresponding entries in the matrices derived by all the different methods from the 12 sets of data described above. The median correlation between the matrix of inferences and the 12 matrices of $\bar{r}_d$ is 0.85, which compares to 0.72 for the correlations of true scores r_T, 0.56 for the correlations of constant deviations r_C and 0.77 for r_{cum}. The differences are all significant by the sign-test. The hypothesis that the correlations among deviations correspond most closely to the structure of inferences is clearly confirmed.

The correspondence between the two structures extends to the level of the correlations. The mean absolute value of the correlations estimated from inferences is 0.271, a relatively low value. This is considerably closer to the mean absolute value of the correlations of deviations (0.194) than to the correlations of true scores (0.590).

---

[22] Results pertinent to the symmetry and linearity of the relations among semantic scales will be reported separately.

The matrix of correlations derived from inferences conforms to the general description provided by Osgood *et al.* The clusters of evaluation and dynamism are clearly present, but the distinction between activity and potency is not clearly maintained. These results agree with the finding of Osgood and Suci who showed that the main factors appear when the *O*s are merely required to note which pole of a scale 'goes with' a particular pole of another scale.[23]

DISCUSSION

The essentials of the present model of the semantic differential can be clarified by showing its similarity to a standard situation in psychophysical research. A group of *O*s are shown a set of familiar objects of different sizes and at different distances. The *O*s provide estimates of the size and the distance of each object. It is easy to create conditions in which cues for size and distance are ambiguous or misleading. Under such conditions, if an unusually small playing card is judged to be of standard size, then its distance will also be overestimated.[24] The size of an unusually large box may be underestimated, and its distance will then be underestimated as well. A positive correlation will be found between the *errors* in judgments of size and distance, regardless of the correlation between the true values actually presented. The latter correlation is entirely determined by *E*, while the former represents *O*'s costribution.

The interpretation of such a psychophysical finding is that *O* has learned that there is a perfect correlation between physical size and objective distance, for any given retinal size. The information is now utilized to make *O*'s judgments on the two dimensions *consistent* with one another. Under some conditions (*e.g.* if cues for distance are ambiguous, while cues for size are determinate) such a process may also improve the validity of judgments.[25]

In the semantic differential, of course, *E*, does not himself determine the true scores of the concepts rated. He determines the correlations among true scores by his method of sampling concepts. He is in the position of asking O to estimate the height and breadth of trunks of trees in a forest, chosen more or less arbitrarily. This analysis suggests that two problems of distinct theoretical significance have often been confounded in correlational investigations of semantic ratings: (1) the study of ecological relationships by sampling of concepts and correlating their true scores; (2) the study of the cognitive system of raters by correlating their errors of judgment. The present study illustrates the need for a separation of these questions, but provides no more than tentative answers. In particular, an extensive comparison of the structure of $\bar{r}_d$ and r_T is certainly required, to proceed beyond the overly general conclusion that the structures are similar. Defining the sub-cultures within which particular true scores have validity is also a problem in need of empirical and theoretical treatment.

---

[23] Osgood and Suci, *op. cit.*, 325-338.

[24] For related demonstrations, see W. H. Ittleson, The constancies in perceptual theory, *Psychol. Rev.*, 58, 1951, 285-294.

[25] For a fuller discussion of the role of processes of inference in enhancing the consistency and validity of cognitive processes, see J. S. Bruner, Going beyond the information given, in J. S. Bruner *et. al.*, *Contemporary Approaches to Cognition*, 1957, 41-70.

One important difference between r_T and $\bar{r}_d$ appears well-established. This is the finding that inferences and correlations of deviations strongly underestimate the correlations of true scores. In particular, it has been shown that the correlations among stable idiosyncrasies in the rating of particular concepts are much lower than are the correlations among true scores, and not substantially higher than the correlations among temporary fluctuations. It is not immediately apparent how this finding can be related to Osgood's interpretation of the semantic differential. On the other hand, it is consonant with the view presented here, in which the true score represents a 'real' value which O estimates in his rating.

In this theoretical context, the relatively low magnitudes of $\bar{r}_d$ and inferences have some implications for an interpretation of the process of inference. A view of cognition which emphasizes inferences is likely to suggest a certain degree of haste, perhaps a calculated taking of risk in trading the security of further observation for the speed of the educated guess. In such a frame of reference, we may well expect people to infer too much. As a matter of fact, they seem to infer systematically too little. It is possible only to speculate that conservativeness of inference may be adjustive in facilitating the apprehension of rare, unexpected events, in an environment where most cues have only limited validity.

SUMMARY

A correlation between attributes of judgments may be due primarily to a real association between the dimensions of the objects of judgment. In order to establish whether O makes any additional contribution, it is necessary to correlate *errors* of judgment. In general, such a correlation

between errors indicates the operation of a process of inference from one attribute to the other.

Ratings given by 60 *O*s to 40 concepts on the semantic differential were analyzed into three orthogonal components: (a) variability among the 'true scores' of concepts; (b) variability among the constant deviations of individual raters; (c) specific deviations of individual rates on particular concepts. A structure of correlations was found among the specific deviation of each individual. Individual differences in the pattern of these correlations are slight. There is little interaction of concept with scales, the same pattern of inferences among attributes being apparently applied to a broad range of concepts. The correlations among scales computed from the *O*s' deviations are closely similar in both pattern and level to the correlations among scales estimated by a group of *O*s. In both cases, the pattern of correlations corresponds to the pattern of association among the true scores of objects, but the *level* of the correlations utilized in inferences is systematically lower than corresponding relationships in the ecology.

A PASSIVE CONDITION FOR RAPID ADAPTATION TO DISPLACED VISUAL DIRECTION

By HANS WALLACH and JEROME H. KRAVITZ, Swarthmore College, and JUDITH LINDAUER, Cedar Crest College

The work to be reported is part of a series of studies devoted to the discovery of examples of perceptual learning rapid enough to make detailed investigation possible. It deals with the adaptation to a displaced visual direction caused by a wedge-prism fixed in front of the eye. With one of its faces in the frontal-parallel plane and with its edge vertical, such a prism displaces the visual field laterally, toward the side of the prism's edge, and thus alters the direction in which objects are seen. With *O*s wearing a 20-diopter prism in front of each eye, Held and Bossom obtained such an adaptation in 1 hr. with 10 out of 15 *O*s; the mean adaptation for the 10 *O*s amounted to 17% of the prismatic displacement of the visual field.[1] The same authors obtained complete adaptation with 8 *O*s by having them wear these prisms for several longer periods, distributed over 2, 3, or 4 days and totaling from 11-21 hr.

Our attempt to obtain reliable and rapid adaptation to displaced visual direction under controlled conditions is based on a scheme which had previously been successful in yielding a rapid modification of a perceptual process.[2] It consists in exposing *O* simultaneously to two different conditions of stimulation which give rise to the same perceptual property and in so altering one of them that it misrepresents the objective situation. A discrepancy is thus produced between the perceptual processes that result from the two conditions of stimulation and may lead to a modification of one or both of these processes. To apply this scheme to the adaptation of visual direction, we paired displaced vision with a directional cue of another kind, the visual vertical, by making a conspicuously represented visual direction coincide with the vertical. This condition was achieved by having *O* tilt his head forward and look down at his legs. The legs then nearly parallel a visual direction while at the same time they visually rep-

---

* Accepted for prior publication July 19, 1963. This work was supported by a grant from the National Science Foundation.

[1] Richard Held and Joseph Bossom, Neonatal deprivation and adult rearrangement: complementary techniques for analyzing plastic sensory-motor coordinations, *J. comp. physiol. Psychol.*, 54, 1961, 33-37.

[2] Hans Wallach and Eileen B. Karsh, The modification of stereoscopic depth-perception and the kinetic depth effect, this JOURNAL, 76, 1963, 429-435.

resent a vertical. When *O* wears a prism, the slant with which his legs are now given will be in conflict with the vertical direction which his legs represent to a standing *O*. As will be reported in detail, an exposure of 10-min. duration to this condition produces a large adaptation to the displaced direction in all our *O*s.

When *O* viewed his legs through a prism whose edge was vertical and to the left, he initially perceived his feet displaced toward the left. His legs appeared to be on a slant, seemingly emerging from his torso at an angle. This apparent oblique position of his legs, however, did not last long; half way through the training period, on the average, *O*s reported that their legs now seemed straight and this state of affairs continued to the end of the 10-min. exposure. It would be a mistake to interpret this outcome as the result of very rapid and complete adaptation to the prismatic displacement. When two different conditions of stimulation conflict with each other, either one may determine the content of experience, they may do so alternately or experience may represent a compromise between the normal outcomes of the resulting perceptual processes.[3] It is much more likely that the change in the apparent position of the legs is due to a shift from the domination of perceptual experience by the directional cues to domination by the cues for verticality. True or not, it should be clear that measurements of the adaptation should be made under conditions that do not involve the vertical as a visual direction.

Two different tests of visual direction were used, a pointing-test and a determination of the apparent forward-direction.

The 'pointing-test' consisted in having *O* point at a visual target under conditions that prevented him from seeing his pointing arm. *O* was asked to stand near a testing table, a horizontal board, 34 × 22 in., which could be raised or lowered. Near the middle of its longer edge and 4 in. above it, was an attachment for a horizontal metal piece on which a mold of *O*'s teeth was fixed. When *O* had the mold between his jaws, the table was at the level of his neck and there was room underneath for his arm to point horizontally. A measuring stick fixed to the far edge of the table was used to read the position of *O*'s pointing finger. Attached to the top of this edge was a rim, 2.5 in. high, which prevented an *O* with long arms from seeing his pointing finger. It also served as a rail for a sliding fixture which carried a socket for a flashlight bulb whose filament was 7.5 in. above the surface of the table. Tests were performed under light or under dark conditions. Under the former condition, the target for the pointing was the edge of a cabinet 16 ft. distant across the room, while, under the latter condition, the target was the glowing filament of the flashlight bulb adjusted to form a clearly visible point of light but too dim to illuminate even its immediate environment. To make this pointing-test as different from the forward-direction test as possible, both targets were located 7° to the left of *O*'s forward-direction when, by means of the mold, his head's orientation was fixed relative to the testing table. The test consisted in one pointing response. *O* raised his arm into a horizontal position in such a way that he felt his finger pointing at the cabinet-edge, or, under dark-room conditions, vertically below the glowing filament. It seems to us that this test was in principle not very different from the test used by Held and Gottlieb, and by Held and Hein, who had *O*s who were unable to see their hands mark points on a plane on which the virtual image of a target-pattern was localized.[4]

---

[3] Hans Wallach and Cynthia M. Norris, Accommodation as a distance-cue, this JOURNAL, 76, 1963, 659-664.

The 'forward-direction' test was used to determine the direction of a target when it appeared to *O* to be straight in front of him. When the test was given in the dark, the same point of light which served in the pointing-test was used. With its socket sliding on the rim of the testing table, the little bulb could be moved horizontally in the frontal-parallel plane of *O* until the glowing point appeared to be straight ahead. In daylight, the target in the forward-direction test was a small black oblong visible on a vertical panel of white plastic which was bent to form a half cylinder, 45 cm. in radius. A horizontal slot 45 cm. long and 0.2 cm. wide was cut into the panel into which the target-piece was inserted and along which it could be shifted from left to right. The panel was simply placed on the testing table. This test was a modification of the egocentric test used by Held and Bossom.[5] They placed *O* on a rotating chair inside a homogeneous drum and asked him so to turn the chair and himself that a target appeared straight ahead of him. Either daylight or dark-room conditions were employed. Wanting to avoid the involvement of vestibular stimulation connected with *O*'s actual rotation, we had the target move while *O* remained stationary.

An individual test consisted here of two settings, one in which *E* moved the target toward the position that *O* judged to be straight ahead coming from the left and the other in which he started the target from the right. The test began with the target clearly too far to one side. After a first judgment was obtained, *E* moved the target to a point closer to the middle, obtained the next judgment, which usually also rated the position of the target as lateral, and so on. When judgments became hesitant, indicating that the critical range was reached, smaller steps were taken. An instruction to move the target back a step was obeyed. Initially we used this informal procedure because we did not know how short-lived a rapidly produced adaptation might be and wished to keep the test as brief as possible; later we wanted to keep the results of different experiments comparable and avoided changing the procedure. Inasmuch as *E* could not read the position of the target in the dark, there was, under dark-room conditions, no danger that he might, through the particular procedure he used, influence the result. Only when *O* judged the glowing filament to be straight ahead was the position of the bulb read with the help of a dim flashlight. Always, the average of *O*'s two settings became his score.

There was a good reason for the use of two so greatly different tests in connection with the same training procedure. If both gave evidence of an adaptation, and, particularly, if they measured adaptation of equal amount, the adaptation would be more likely to consist in the modification of a single *perceptual* process and the possibility of its connection with a particular testing *response* would be distinctly lessened.

---

[4] Richard Held and Norman Gottlieb, Technique for studying adaptation to disarranged hand-eye coördination, *Percept. Motor Skills*, 8, 1958, 83-86; Richard Held and Alan V. Hein, Adaptation of disarranged hand-eye coordination contingent upon re-afferent stimulation, *ibid.*, 8, 1958, 87-90.

[5] Held and Bossom, *op. cit.*, 33-37.

Each test could be given under two conditions: while *O* wore the prism (correction-method) or with normal viewing (alteration-method). With the edge of the prism to the left, all objects appeared displaced to the left, initially by approximately 11°. It will be seen that a pointing-test measured this effect of the prism with fair accuracy; on the average *O*s pointed about 11° to the left of the target. After some adaptation had taken place, pointing was somewhat to the right of the initial direction, that is, pointing was shifted toward the true direction of the target, in partial compensation for the effect of the prism. Adaptation was, therefore, measured as a correction of the error which the prism caused—hence the name we have given to this method. Since our prism displaced all visual directions to the left, an object which was seen through the prism as straight in front was actually approximately 11° to the right. Thus, in the initial forward-direction test, the glowing filament set to appear straight ahead was actually 11° to the right. After some adaptation had taken place, the filament had to be set less than 11° to the right in order to appear straight ahead. Again, adaptation was measured as a reduction of the prism-error.

When the pointing-test was made *without* the prism, *O* pointed correctly at first, while adaptation made itself felt as a pointing error: *O* pointed to the right of the target. Since adaptation makes objects appear displaced to the right, the glowing filament viewed without prism was actually to the left of the true forward-direction when *O* perceived it as straight ahead. Without the prism, adaptation was measured on both tests as an alteration of the initial veridical measurement.

In the earlier work mentioned, adaptation was measured by what we call the 'alteration-method.' We used the correction-method in addition on the chance that it might yield a greater adaptation-effect. In his work with long adaptation-periods, Ivo Kohler discovered conditional adaptation: long after reversing glasses had been omitted and all adaptation-effects had disappeared, wearing the glasses' frames would cause some of the effects to occur again.[6] If, in our experiments, adaptation were partially conditional, the correction-method, where the prism was worn for the tests, should yield larger effects. We were concerned with this possibility, because prism-wearing is associated not only with the tactile sensations and the narrowing of the visual field produced by the goggles, but also with specific form-distortions such as the curving of contours parallel to prism's edge. In our experiment, there was also a chromatic filter which put a strong green color across the visual field.

A prism of 20 diopters produces rather strong colored fringes along contours, and thus causes a serious loss in acuity. We eliminated these fringes and greatly improved acuity by using, along with the prism, a Wratten Filter 58 by Kodak which passes a fairly narrow band of frequencies near the center of the visible spectrum. A plastic prism, 4.8 cm. from edge to base and 3.7 cm. wide, and the filter were attached to welder's goggles from which the glasses had been removed. The goggles were slightly altered to bring the inner prism-surface into frontal-parallel orientation. The prism was always in front of the right eye with its edge to the left. *O*'s left eye was always occluded by having the opening of the goggles for that eye covered and by having him wear an eye-patch when the goggles were not used.

---

[6] Ivo Kohler, Über Aufbau und Wandlungen der Wahrnehmungswelt, *SB Öst. Akad. Wiss.*, 227, 1951, 1-118.

Unless otherwise stated, adaptation was produced by having *O* stand relaxed with head bent forward looking down at his lower legs. *O*s had no trouble holding this position for 10 min., the length of the training period in the present series of experiments. At the start of each experiment, an impression of *O's* teeth was taken and fixed to the testing table. The table's height was so adjusted that the mold fitted *O's* mouth comfortably when he was standing straight at the testing table. *O* was instructed to hold the mold firmly between his jaws during all tests.

For easier reference the following abbreviations are introduced: $P =$ pointing-test; $F =$ forward-direction test; $c =$ correction-method, test with prism; $a =$ alteration method, test without prism; $1 =$ test taken before training; $2 =$ test taken after training; $Tr =$ 10-min. training period. Thus, *Pa*1 means a pointing-test taken without prism prior to training.

EXPERIMENT I

The first experiment to be reported—an improvement on a number of preliminary pointing experiments—was done under daylight conditions. *O* was given a forward-direction test and a pointing-test, both without prism. Then the goggles were put on him and a pointing-test and a forward-direction test were given, in that order. There followed the training period and the four post-training tests, which were given in this order: *Fc*2, *Pc*2, *Pa*2, and *Fa*2. Four pairs of measures of a potential adaptation-effect were thus made, employing each test with both methods. In all, 15 *O*s participated. It is an essential feature of this sequence that the tests with prism immediately precede and follow the training period, and that *O* put on the goggles only once. Removing the prism between the pre- and the post-training tests might introduce an error, because the displacement-angle varies somewhat with the angle that the line of sight forms with the prism. A deviation of the line of sight toward the prism's edge will enhance the angle of displacement, while the opposite deviation will decrease it.

Results of the pointing-tests. The adaptation to the prismatic displacement is measured by the mean difference between the pointing directions before and after training. Tested with prism, this difference amounted to 4.55°, and was significant beyond the 1% level ($t = 9.12$). It can be compared with the displacement-angle which the prism causes for the particular target-direction, which is computed as the mean difference in the pointing directions obtained prior to training without and with prism ($Pa1 - Pc1$). It amounted to 11.82°. Thus, adaptation in this group of *O*s was 39.3% of the initial prismatic displacement.

Measured by the alteration-method, the difference between the pointing-directions before and after training was somewhat smaller; namely, 3.88° ($t = 5.04$, $p < 0.01$). As mentioned earlier, a smaller training effect measured without the goggles could be explained as the result of a partial conditional adaptation. Although we obtained a corresponding difference of similar magnitude in our next experiment, it did not reach significance. In an analysis that combined the results of Experiments I and II, the difference in the adaptation-effects measured with and without prisms was not quite significant ($F = 2.136$, $df = 1, 28$).

The forward-direction tests did not succeed. The mean directions in which the target appeared straight in front did not differ appreciably before and after training. As the next experiment shows, this outcome was due to some characteristic of the daylight-conditions under which the tests were performed, and not to the nature of the test as such.

EXPERIMENT II

This experiment was a repetition of the preceding one, except that all tests were performed in the dark, with the glowing filament serving as target, and that a small change was introduced in *O*'s posture during the training period. While in Experiment I, *O* was asked not to move his head after he had tilted it forward and started to look at his legs, no such instructions were given in Experiment II. Instead, changes in the position of *O*'s head were recorded by means of a small mirror attached to the head which reflected light from a projector upon a 5-ft. vertical scale attached to the wall. A reading of the position of the spot of reflected light was taken every 30 sec. during the training period.

The *results of the pointing-tests* were quite similar to those of Experiment I. The adaptation-effect measured with prisms was 4.15° and was significant beyond the 1% level ($t = 4.19$). With the initial prismatic displacement ($Pa1 - Pc1$) measured here as 11.52°, adaptation amounted to 37.4% of that displacement. When the test was done without the prism, the effect was again somewhat smaller; namely, 3.76° but highly significant ($t = 8.68$, $p < 0.01$).

The *forward-direction test* in the dark was completely successful. Moreover, it measured an adaptation of the same amount as the pointing-test. With the prism, the mean difference between the scores taken before and after the training period was 3.76° ($t = 4.0$, $p < 0.01$). While this value was slightly smaller than that obtained with the corresponding pointing test (4.15°), the mean change in the forward-direction due to the prismatic displacement ($Fa1 - Fc1$) also was slightly smaller; namely, 9.63°; expressed as a fraction of the initial prismatic displacement, adaptation amounted to 39.1%. About one half of this decrement in the measured prismatic displacement was due to the fact that, in the forward-direction test, the line of sight formed a different angle with the prisms than it did with the target for the pointing-test. Once more, the test without prism showed a smaller adaptation-effect (3.18°, $t = 3.57$, $p < 0.01$) than the one with prism, and again the difference was not significant.

Very slow changes in the position of *O*'s head occurred during the training period when *O* had no instructions to keep the head still. They were found to be unrelated to the adaptation-effect. An average deviation of the 20 head-position readings taken during training was computed for each *O* and rank-order correlations were obtained between these measures and the adaptation-effects measured with the *Pc* and the *Fc* tests. No correlations were found. For the *Pc* test, $r' = -0.04$ ($N = 19$), and, for the *Fc* test, $r' = -0.08$ ($N = 12$).

Since the results of the forward-direction tests are in good agreement with those of the pointing-tests when they were performed in the dark, the question of why the forward-direction test in daylight was not successful becomes an issue. There are facts which suggest an explanation. The variability of the initial forward-direction scores was considerably smaller under daylight-conditions than in the dark; the *SD* for the former was 1.39°, while for the latter it was 3.46°, and this difference was significant at the 1% level. (No such difference between the conditions of daylight and darkness was found for the pointing-tests; the corresponding *SD* were 3.25° and 2.72°.) The variability of judgments within individual *O*s also was different under the two conditions. Since an *O*'s forward-direction score was an average of two settings, judgments made by individual *O*s could be compared; they turned out to be more consistent under daylight conditions. The mean difference between the two settings in the initial forward-direction test amounted to 1.50° for daylight and 4.22° for darkness. There was probably a tendency operating to center the target in the visual field. This explanation is supported by the fact that the variability in the scores for the first forward-direction test *with* prism was quite high ($SD = 4.54°$). When the prism was worn, some content of the visual field probably was so displaced that its center did not coincide with the apparent forward-direction, and such a discrepancy could account for the enhanced variability. This interpretation is supported by the fact that in daylight the forward-direction test did not measure the prismatic displacement correctly. The difference between the mean settings with and without the prism amounted to 6.15°, only 63% of the prismatic displacement measured in darkness.

EXPERIMENTS III A, B, AND C: WITH THE ALTERATION-METHOD

In the preceding experiments, for the reason already stated, the tests without prism always followed the tests by the correction-method. If the adaptation that is produced in our brief training period dissipates rapidly, these experiments do not give a correct picture of the adequacy of the alteration-method. For this reason, we did three experiments in which a test without prism was made immediately following the training period, and the same test was repeated a number of times to check on the course of possible dissipation. Where the pointing-test was used, *O* made four pointing responses before and eight after training. In the case of the forward-direction test, *O* made the same number of individual settings, which means that two scores were obtained before training and four afterwards. The three experiments were done under the following conditions: Experiment III A, pointing-test in daylight; Experiment III B, pointing-test in darkness; Experiment III C, forward-direction test in darkness.

TABLE I

MEAN ADAPTATION-EFFECTS OBTAINED IN EXPERIMENT III A, B, AND C, COMPARED WITH THE 2 PREVIOUS EXPERIMENTS

($N=15$)

Averages	Pointing, daylight		Pointing, darkness		Forward-direction, darkness	
	Mean	*t*	Mean	*t*	Mean	*t*
All responses	3.55°	6.65	4.19°	11.21	3.15	2.94
First responses	4.18°	6.15	3.81°	7.56	2.39	2.42
Experiments I and II	3.88°	5.04	3.76	8.68	3.18	3.57

Table I shows the adaptation-effects obtained. They are computed in two ways: from the average of all the responses made before training and the average of all responses made after training (first row) and from the first response or score before and the first after training only (second row). The latter method corresponds to the one used in Experiments I and II, and the corresponding measurements from these experiments also are given (third row). The reasons for the differences in the results of the two methods of computation become clear from an inspection of Fig. 1

where the results of the individual responses are represented. The plots also show that there was no appreciable decline in the adaptation-effect in the course of the repeated tests. A comparison of the means in the second and the last row of Table I show that the delay of the tests without prism in Experiments I and II did not cause a lowering of the adaptation-effect.

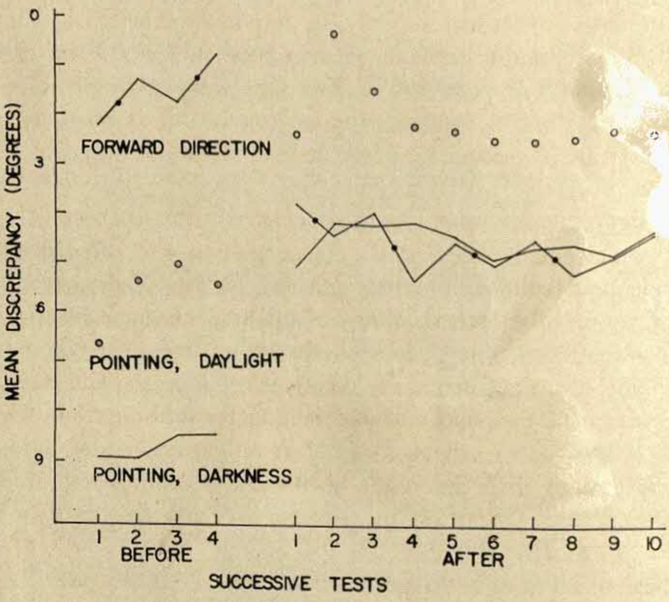


FIG. 1. RESULTS OF EXPERIMENT III

The mean effects were not appreciably greater when these tests followed the training period immediately. The small and insignificant difference in the third column ($t = 0.61$) was in the opposite direction.

EXPERIMENT IV

Our final experiment was concerned with the cause of the adaptation. When *O* regards his legs through the prism, he is, as has been stated, exposed to conflicting cues. His legs nearly parallel a visual direction and they also represent the vertical direction. With the visual direction displaced by the prism, there is a discrepancy between cues representing the visual direction and the cues on which the perception of the vertical direction is based. This discrepancy can be eliminated by having *O* lie horizontally on his back, with his head raised to permit him to see his legs through the prism. Would adaptation occur under these conditions? We thought not. There would remain only a discrepancy between the displaced visual location of the legs and their location according to postural data. Results by Held and Hein seemed to indicate that such a discrepancy would be insufficient to cause an adaptation.

During the training period, *O* was lying flat on a day-bed, his head propped up by a firm cushion. While in the case of the standing *O* the prismatic displacement led, initially, to a vivid experience, the supine *O*s hardly noticed the displacement of their legs; in fact, 12 out of the 23 *O*s participating never reported it. Two tests were employed: a pointing-test in daylight and a forward-direction test in the darkness, both done with the prism. The sequence of events was the following: *Fc*1, *Pc*1, *Tr*, *Pc*2, *Fc*2.

Pointing responses were very little changed after training. The mean difference amounted only to 1.16° (compared to 4.55° in Experiment I under identical testing conditions) and was, in spite of the larger number of *O*s (23) and the high reliability of this test, insignificant ($t = 1.50$). The forward-direction test, on the other hand, measured an adaptation much higher than had previously occurred in our work. The mean difference between the pre- and the post-training forward-direction was 6.22° ($t = 9.69$, $p < 0.01$), about 65% of complete adaptation. It was significantly greater than the result of the corresponding test in Exp. II (where the mean change in the forward-direction amounted to 3.76°) with $t = 2.19$ and $p < 0.05$.

Because of its unexpected result, Experiment IV was repeated with the prism in reversed orientation. While in the experiments so far reported the edge of the prism was always to the left, the edge was now to the right and the prism displaced the visual field to the right. Under these conditions, the forward-direction test measured a similarly high adaptation, with the mean effect amounting to 5.46° (15 *O*s).

DISCUSSION

One of the distinctive features of our experiments lies in the fact that two kinds of tests followed the same novel training procedure. One test required a physical response (pointing-test) and the other a perceptual judgment (forward-direction test). This confrontation of two kinds of test greatly different in nature yielded the most interesting results of our work. Previously, physical response-tests had been used in connection with training conditions which involved visually observed active movements of an arm,[7] and judgments of perceived direction in connection with *O*'s active locomotion.[8]

---

[7] Held and Gottlieb, *op. cit.*, 83-86; Held and Hein, *op. cit.*, 87-90. Recent work

In our Experiment II, the two kinds of test measured adaptation in the same amount, and this result seemed to support the notion that the two tests were concerned with different manifestations of *one* perceptual-learning effect, an altered evaluation of the given visual direction; however, Experiment IV, in which *O* observed his legs through the prism while lying horizontally, contradicted this notion. While the forward-direction test showed a very high adaptation-effect, the yield of the pointing-test was very small and insignificant, the difference in the results of the two tests being very reliable ($t = 4.34$, $p < 0.01$). Having *O* lie down during training affected the results of the two tests so decisively that we must conclude that the tests measure the effects of different processes. In the light of Experiment IV, the fact that the two tests gave identical results in Experiment II becomes, in turn, a problem. A change in *O*'s position during training from supine to standing not only caused the pointing test to yield positive results, but also significantly *de*creased the adaptation measured with the forward-direction test. An explanation of this peculiar combination of facts must be left to further study. Nor do we venture any statement concerning the nature of the two adaptation-processes which seem to be responsible for our results.

The other important result of our work has to do with the passive condition of our *O*s during training. Held and his collaborators maintain that adaptation to displaced visual direction can be obtained only when *O* visually observes the effects of his own active movements, and they have consistently obtained results supporting this view, with a test involving judgment of perceived direction as well as with a physical-response test. In our experiments, adaptation of high degree was obtained with both kinds of test under conditions in which *O* remained passive. A particularly striking result occurred in Experiment IV, where *O*s were supine and did not even have to maintain their equilibrium.

---

by Charles S. Harris (Adaptation to displaced vision: visual, motor or proprioceptive change? *Science*, 140, 1963, 812-813) also belongs here. The explanation Harris offers for his results—that adaptation consists in "a change in the felt position of the arm relative to the body"—is clearly not applicable to our results. During training our *O*s neither moved nor saw their arms.

[8] Held and Bossom, *op. cit.*, 33-37.

EXPERIMENTALLY PRODUCED MEANINGFULNESS (m) IN PAIRED-ASSOCIATE LEARNING

By GEORGE V. C. PARKER and CLYDE E. NOBLE,
Montana State University

Since the introduction of Noble's scale of *meaningfulness* (m),[1] investigators have consistently found a positive relationship between rate of acquisition and m in verbal-learning tasks. The role of meaningfulness has been studied most analytically by varying independently the m-values of the S- and R-terms in paired-associate learning. Confirming a prediction made by Noble and McNeely,[2] the data typically show that the major influence of m is on the R-term.[3] No one has yet attempted the experimental alteration of m by teaching *S*s new associations to items of low normative m-value, then testing for a positive transfer-effect in paired-associate learning when these items are used as R-terms. The present paper describes two investigations of this problem.

In view of the high correlation between m and judgments of *familiarity* (f), care must be exercised in the interpretation of any study reporting positive results. Could the influence of m derive from differential familiarization? Since high m implies high f by definition, its *modus operandi* might be sheer frequency of experience (n). Although reasonable, this hypothesis is inconsistent with three facts: (1) the regression of f on m is nonlinear;[4] (2) serial learning but not associative value is facilitated by n;[5] and (3) the S-term is the only locus of an effect of n in paired-asso-

---

[*] Received for publication October 2, 1961. This work was supported in part by a contract between the Office of Naval Research and Montana State University. Helpful comments were received from Drs. Robert B. Ammons, Arthur W. Melton, and Charles D. Parker. Mr. G. V. C. Parker is now at The State University of Iowa, Iowa City, Iowa.

[1] C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421-430.

[2] C. E. Noble and D. A. McNeely, The role of meaningfulness (m) in paired-associate verbal learning, *J. exp. Psychol.*, 53, 1957, 16-22.

[3] V. J. Cieutat, F. E. Stockwell, and C. E. Noble, The interaction of ability and amount of practice with stimulus and response meaningfulness (m, m') in paired-associate learning, *ibid.*, 56, 1958, 193-202; V. J. Cieutat, Supplementary report: Stimulus and response meaningfulness (m') in paired-associate learning by hospitalized mental patients, *ibid.*, 58, 1959, 490; R. G. Hunt, Meaningfulness and articulation of stimulus and response in paired-associate learning and stimulus recall, *ibid.*, 57, 1959, 262-267; B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Learning*, 1960, 27-43.

[4] Noble, The meaning-familiarity relationship, *Psychol. Rev.*, 60, 1953, 89-98.

[5] Noble, the effect of familiarization upon serial verbal learning, *J. exp. Psychol.*, 49, 1955, 333-338; D. A. Riley and L. W. Phillips. The effect of syllable familiarization upon rate of learning, association value, and reminiscence, *ibid.*, 57, 1959, 372-379.

ciate learning when the present methods and materials are used.[6] The following two experiments were designed to test the hypothesis that number of associations rather than familiarization is the primary source of the effect of meaningfulness.[7]

EXPERIMENT I

The purpose of Experiment I was to have *S*s first learn different numbers of arbitrary associations to low-*m* items, test for recall, then measure transfer-effects in a paired-associate task. There were three phases of the experiment: Phase 1 (Training), Phase 2 (Recall), and Phase 3 (Transfer). Placing items with experimentally endowed *m* in the *response*-positions, where *m* exerts its strongest influence, allowed two predictions: (1) that level of acquisition should be greater for those pairs containing higher acquired *m* than for pairs with lower original normative *m*; and (2) that the functional relationship between performance and experimentally produced *m* should be similar to that discovered in studies of normatively defined *m*.

Apparatus. Two Stoelting memory-drums adjusted to a 2-sec. rate of presentation were used for all parts of the experiment. The verbal materials were pronounceable dissyllables selected from norms recently developed for the Montana population.[8] This new *m*-scale has a correlation of 0.97 with the original 1952-scale but the values are higher by an average of 2.11 units. All items were typed in black capital letters on white tapes.

Subjects. The *S*s were 20 volunteers from classes in psychology at Montana State University during 1958-59, all naïve to paired-associate learning. Each *S* was assigned to one of four list-conditions as he reported to the laboratory.

Procedure. Phase 1 (Training) involved presentation of lists containing three items paired differentially with three, six, or nine associations. Four such lists were prepared having two sets of S-terms (Set A, Set B) paired with two sets of associations. For Set A the S-terms were LATUK, NARES, VOLVAP; for Set B the S-terms were ULNA, BALAP, SAGROLE. The R-terms were so selected that the median *m*-value ($\dot{m}$) for all subsets was nearly the same (range = 7.79 to 8.30), but one of each pair of subsets was composed of items from the center of the *m*-scale while the other pair came from the scale's extremes. The S terms were all of low *m*-value (range = 2.64 to 3.60). Ten judges checked the lists for associations which might either hinder or facilitate learning. The *S* was instructed as follows:

---

[6] D. R. Gannon and C. E. Noble, Familiarization (*n*) as a stimulus factor in paired-associate verbal learning, *ibid.*, 62, 1961, 14-23; D. T. Hakes, The role of stimulus and response familiarization (*n*) in paired-associate verbal learning, paper reported to the Midwestern Psychological Association, Chicago, Ill., 1961.

[7] Noble, Meaningfulness and familiarity, in C. N. Cofer and B. S. Musgrave (eds.), *Verbal Behavior and Learning: Problems and Procedures*, 1963, 76-119.

[8] C. E. Noble and G. V. C. Parker, The Montana scale of meaningfulness (*m*), *Psychol. Rep.*, 7, 1960, 325-331.

When we begin, you will see two words through the window, *here*. In all cases, the word on the *right* will be one of the *meanings* to be learned for the word on the *left*. You are to pronounce both of these when they appear. They will be followed by another pair of words, and you will pronounce these also. Continue pronouncing each pair of words as they appear in the window. Do you have any questions? After several trials you will be tested to see how well you have learned the meanings, so do the very best you can to remember *all* the meanings for each word.

All *S*s were given 20 trials with a 4-sec. interval between trials to learn 3 associations to one item, 6 associations to another item, and 9 associations to still another item. The sequence of pairs was reversed on alternate trials. After adding the mean *preëxperimental* or normative *m̈*-values, the conditions adding 0, 3, 6, and 9 associations had respective net *m*-values of 3.06 (low), 6.18 (below average), 9.13 (above average), and 11.67 (high) associations.

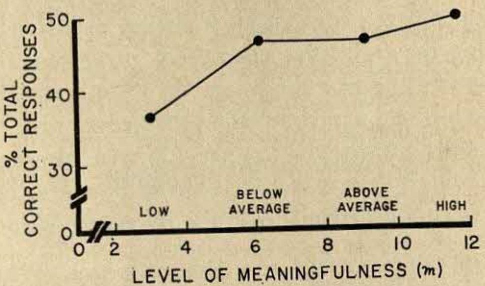


FIG. 1. PERCENTAGE TOTAL CORRECT RESPONSES IN 20 TRIALS AS A FUNCTION OF LEVEL OF INDUCED MEANINGFULNESS IN EXPERIMENT I

Phase 2 (Recall) consisted of 2-min. test-periods at the end of Trials 8, 12, 16, and 20 of Phase 1. During this period, *S*s were presented the three S-items alone on a mimeographed form and required to write under each their respective associations. Failure to reach the criterion of 100% recall by the end of Trial 20 resulted in rejection of *S*.

Phase 3 (Transfer) was a paired-associate learning situation in which *S*s were given 20 trials with four randomized lists of eight S-R pairs. Standard instructions were used,[9] and both S and R were pronounced. Three R-terms were the S-terms from Phase 1, and five others were new ones of similar *original* *m̈*-value. The new S-terms were all of low *m̈*-value (range = 1.79 to 4.92) and the intertrial-interval was 12 sec.

Results and discussion. The results of Experiment I offer support for the two predictions which generated it. Level of acquisition was greater for pairs containing higher acquired *m* than for pairs of lower original normative *m*, and the functional relationship between performance and experimentally produced *m* appeared to be similar to that discovered with normatively defined *m*. The general results are summarized in Fig. 1, showing the percentage of correct responses ($R\%$) at each level of m during 20 trials (N). The relationship is positive and suggestive of negative acceleration. Since a monotonic function was expected, a correlated t-test between the mean correct responses ($R+$) at the low ($m = 3.06$) and high ($m = 11.67$) levels of meaningfulness is appropriate.[10] For 19 df the ratio of 5.29 is significant ($P < 0.01$). We conclude that there is a tendency for amount of facilitation in Phase 3 to increase with the number of associations acquired in Phase 1.

---

[9] Gannon and Noble, *op. cit.*, 17.

EXPERIMENT II

Although Experiment I suggested that number of associations is a major factor producing the facilitative effective of m, it contained no control for a familiarization effect on the R-term. In Experiment II, therefore, the influence of both n and experimentally produced m were investigated together, testing three predictions: (1) and (2) were identical with those in Experiment I, while (3) was that there would be no facilitative influence as a result of having differential n in the response-position.

Apparatus. Equipment and timing were identical to those in Experiment I.

Subjects. The *Ss* were 144 volunteers from classes in psychology at Montana State University during 1959-60, all naïve to paired-associate learning. Ages ranged from 17 to 53, with a mean of 20.5 yr. There were 89 men and 55 women, assigned to the treatments as they reported to the laboratory. Twenty additional *Ss* had to be used (in Groups E and E-F) because of failure to meet the criterion of recall in Phase 2.

Procedure. There were four treatment-groups symbolized as follows: Experimental (E), Control (C), Experimental-Familiarized (E-F), and Control-Familiarized (C-F). Groups E and C contained 48 *Ss* each, while Groups E-F and C-F contained 24. The conditions were usually assigned to *Ss* in the order: 2 Es, 2 Cs, 1 E-F, and 1 C-F.

Group E was treated like the *Ss* in Experiment I, except for some changes in the S-terms of Phase 1 which reduced the variance of their $\dot{m}$-values (range $= 2.69$ to 2.83). The design of the Training Phase for Groups E and E-F is outlined in Table I. After Phase 2 (Recall) Group E received Phase 3 (Transfer), consisting of 20 trials on an eight-pair list in four random orders as in Experiment I. Treatment of *Ss* in Group E-F was identical, except that they were familiarized with five items to appear as R-terms in Phase 3 but not practiced in Phase 1: the two control terms GOJEY and QUIPSON plus either experimental Set A (BRUGEN, ZUMAP, SAGROLE) or experimental Set B (ULNA, NARES, VOLVAP). These five items were presented in a random sequence at a 2-sec. rate, and no item appeared more than twice consecutively. The *Ss* in Group E-F pronounced each item 16 times before Phase 1 and again after Phase 2 for a total of 32 exposures (n) before its appearance as an R-term in Phase 3. Experiments on familiarization, using these materials, have shown that facilitative returns tend to be asymptotic around $n = 20$,[11] but we used 32 exposures for good measure.

---

[10] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 342.

Group C received only the treatment of Phase 3. Items of high normative $\bar{m}$-value were substituted for those items to which Groups E and E-F had learned extra associations. Two sets of words (corresponding to Set A and B of Phase 1) were selected so that mean normative $\bar{m}$-values were matched as closely as possible with net experimental m-values, *i.e.* SEQUENCE and QUOTA (5.76) vs. BRUGEN and ULNA (2.80 + 3), TYPHOON and VILLAGE (8.72) vs. ZUMAP and NARES (2.77 + 6), MONEY and KITCHEN (11.12) vs. SAGROLE and VOLVAP (2.67 + 9). The means of the four m-levels (rounded to one decimal place) will be referred to as low (2.8), below average (5.8), above average (8.7), and high (11.4). Group C-F received only the treatment of Phase 3, preceded by familiarization on GOJEY and QUIPSON (plus either Set A or B) as with Group E-F. For the Transfer-phase Group C-F used the same randomized lists as Group C. These *Ss* pronounced each of the 5 low-$\bar{m}$ items 16 times, were engaged informally in conversation for 5 min. (replacing Phases 1 and 2), pronounced the items 16 more times, and immediately moved to the Transfer-phase. The paired-associate lists of Phase 3 are reproduced in Table II.

TABLE I

DESIGN OF THE TRAINING PHASE OF EXPERIMENT II FOR GROUPS E AND E–F

Condition (20 trials each)	Stimulus-terms		Mean original m-value	Added Associations	Net m-value
	Set A	Set B			
3	BRUGEN	ULNA	2.80	MALLET KITCHEN QUOTA	5.80
6	ZUMAP	NARES	2.77	KEEPER JELLY HEAVEN EFFORT WAGON YEOMAN	8.77
9	SAGROLE	VOLVAP	2.67	YOUNGSTER REGION ZERO LEADER TYPHOON INCOME UNCLE FATIGUE VILLAGE	11.67

Results. The general results of Experiment II are summarized in Fig.

---

[11] Noble, *op. cit.*, 1955; Gannon and Noble, *op. cit.;* Hakes, *op. cit.;* R. W. Schulz and I. F. Tucker, Stimulus familiarization in paired-associate learning, paper reported to the Psychonomic Society, New York, N.Y., 1961; Noble, *op. cit.*, 1963.

2, showing the over-all proficiency of the four groups at each level of m during Trials 1–20. The m-values marked along the abscissa represent averages of the normative and experimental values as explained above. A 4 × 4 Type I analysis of variance,[12] performed on the numbers of correct responses ($R+$), revealed significant main effects of Meaningfulness ($F = 50.98$, $df = 3/420$, $P < 0.001$), Conditions ($F = 4.11$, $df = 3/140$; $P < 0.01$), and their interaction ($F = 8.61$, $df = 9/420$, $P < 0.001$). It is evident from Fig. 2 that correct performance is a positive function of increasing m, although the influence of *normative m* (open data-points) is greater than that of *experimental m* (filled data points) at higher levels. Comparison of Fig. 1 and Fig. 2 shows that the rising trend of Experiment I is essentially replicated by Group E in Experiment II.

TABLE II

LIST USED IN THE TRANSFER-PHASE OF EXPERIMENT II FOR ALL FOUR GROUPS

Groups E and E–F	Groups C and C-F	
	Set A	Set B
FERRULE-VOLVAP	FERRULE-VOLVAP	FERRULE-KITCHEN
GOKEM-ULNA	GOKEM-ULNA	GOKEM-QUOTA
KAYSEN-SAGROLE	KAYSEN-MONEY	KAYSEN-SAGROLE
LEMUR-QUIPSON	LEMUR-QUIPSON	LEMUR-QUIPSON
MEARDON-GOJEY	MEARDON-GOJEY	MEARDON-GOJEY
NOSTAW-BRUGEN	NOSTAW-SEQUENCE	NOSTAW-BRUGEN
POLEF-NARES	POLEF-NARES	POLEF-VILLAGE
WELKIN-ZUMAP	WELKIN-TYPHOON	WELKIN-ZUMAP

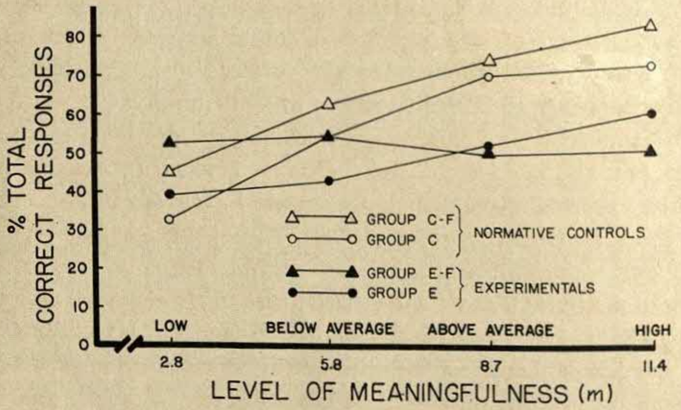


FIG. 2. PERCENTAGE TOTAL CORRECT RESPONSES IN 20 TRIALS FOR VARIOUS CONDITIONS AS A FUNCTION OF LEVEL OF MEANINGFULNESS IN EXPERIMENT II

---

[12] Lindquist, *op. cit.*, 266-273.

A striking feature of Fig. 2 is the curve for Group E-F, which is nearly parallel to the abscissa. By a trend-analysis,[13] the slope of this line is essentially zero ($F < 1$) and is significantly different from each of the other three ($F \geq 3.78$, $df = 3/420$, $P < 0.01$). With respect to vertical comparisons, Groups E and E-F differ significantly only on the low *m*-items ($t = 2.81$, $df = 70$, $P < 0.01$). Hence, although the trends for Groups E and E-F are dissimilar, their scores are not significantly different at levels where *m* was built in. Focusing now on the performance of all groups with the low items, it can be seen that Groups E-F and C-F (triangles) both performed better than their respective non-familiarized counterparts (circles), the difference between Groups C and C-F at this level giving a *t* of 2.63 ($df = 70$, $P < 0.01$). Thus there was an apparent incremental effect of familiarization upon performance with the items of low-*m* value.

Over-all practice curves, disregarding *m*-levels, were next subjected to a 5×4 Type I analysis of variance: five blocks of Practice by four Conditions. There were significant main effects of Practice ($F = 408.47$, $df = 4/468$, $P < 0.001$) and Conditions ($F = 3.80$, $df = 2/117$, $P < 0.05$), but no interaction ($F = 1.50$, $df = 8/468$, $P > 0.20$). Levels of proficiency were generally higher for normative-*m* items, and there was a consistent tendency for performance to be superior with the added influence of familiarization. This suggests that the procedure we used to produce *m* failed to build in all the facilitative attributes present in higher normative *m*-items, even though the trends in performance are similar. Were it not for the experiment of Gannon and Noble and its replication by Hakes,[14] neither of which showed any facilitation accompanying familiarization of the R-term, these data would also be grounds for reconsidering the hypothesis of Cieutat, Stockwell, and Noble, and of Underwood and Schulz that *n* may be adequate to account for the effect of *m*.[15] The issue is more complicated than this, however, as can be seen in the following four-part analysis by different levels of *m*.

The curves of acquisition of the four treatment-groups at each level of *m* over five 4-trial blocks of practice are presented in Fig. 3. The familiarized groups (triangles) are more proficient at the lowest level of *m*, but as *m* increases the control or normative groups (open points) become superior. Fig. 3 also shows that the slopes of the curves of acquisition vary directly with *m*, as found earlier.[16] The results of of four 5 × 4 Type I analyses of variance performed on these data indicated that the main effects of Practice ($F \geq 191.55$, $df = 4/560$) and Conditions ($F \geq 3.16$, $df = 3/140$) are significant at all levels of *m* ($P < 0.025$), but the Practice × Conditions interactions ($F \geq 1.93$, $df = 12/560$) are significant only at the low ($P < 0.05$) and above average ($P < 0.001$) levels of meaningfulness. Evidence of changing orders and non-parallel trends requires a cautious interpretation of the hypothesis of familiarization. Any statement implying a general superiority of the familiarized groups in Experiment II would be misleading; Group C-F is consistently above Group C, but Group E-F drops below Group E at high *m*.

---

[13] Lindquist, *op. cit.*, 350-354.

[14] Gannon and Noble, *op. cit.*; Hakes, *op. cit.*

[15] Cieutat, Stockwell, and Noble, *op. cit.*; Underwood and Schulz, *op. cit.*, 84-129.

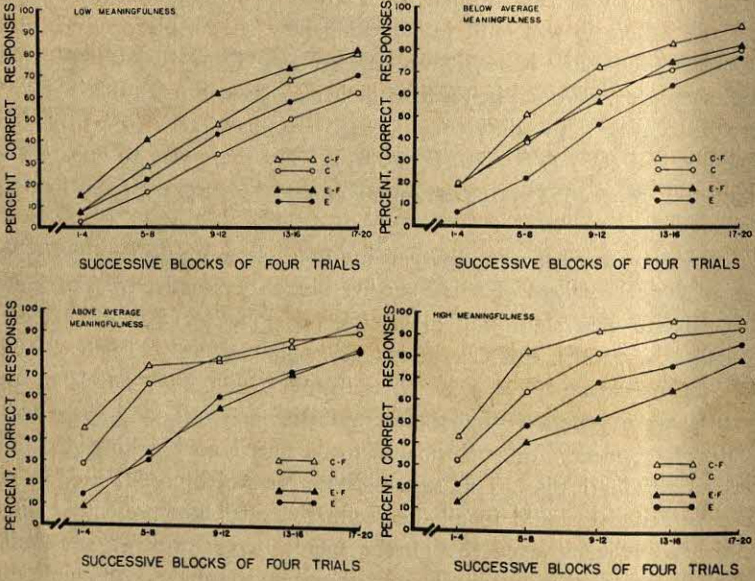


FIG. 3. PERCENTAGE CORRECT RESPONSES FOR VARIOUS CONDITIONS IN EXPERIMENT II AS A FUNCTION OF AMOUNT OF PRACTICE, WITH THE FOUR LEVELS OF MEANINGFULNESS ANALYZED SEPARATELY

Discussion. The implications of Experiments I and II are that *m* can be produced by experimental manipulation, and that it has the classical facilitative effect on performance. Comparison of learning curves of *S*s practicing with items of equivalent values of normative and laboratory-produced m reveals a close similarity. The proficiency level for experimental-m items is usually lower than that for corresponding normative-m terms, but both curves indicate that learning is a positive function of m, defined either way. One reason for the difference is that items of higher normative m-value probably have more associations connected to them than those which *Ss* write during a brief interval of time. Building up m to a given level by having *Ss* learn only enough associations to fill in the differential between low-m and high-m items may result in fewer associations than would be required to produce the facilitative effect of higher normative m. Responses to high-m items generally appear to be of high m themselves and seem to have a relationship to each other. For example, KITCHEN leads to such responses as *food, mother, hungry,* and *table*—all of which belong together and which tend to elicit common associations. Since Phase 1 employed arbitrary responses with inter-set and intra-set associations minimized, the present method may have retarded the establishment of new S-R bonds. Many *Ss* reported an inability to learn that 'S *means* R' because the responses were so "mixed-up," suggesting that they tried to learn the R-set as a cluster under the heading of the S-term. A question is thus raised as to the necessary and sufficient conditions responsible for the facilitative effect of m. Aside from merely producing multiple associations to stimuli, must they also 'belong' and be 'meaningful' themselves? Essentially, this question reduces to the more fundamental issue of whether meaningfulness is a unitary or multivariate attribute.[17]

---

[16] Noble and McNeely, *op. cit.*, 18.

Turning now to the role of n, it appears that the procedure of familiarization had a facilitative influence at the low-m level but a slight decremental effect on the unfamiliarized, higher-m terms (Group E-F). Practice in articulating low-m responses is important, for some nonfamiliarized *Ss* (Groups E and C) were still having difficulty pronouncing the paralogs on the last trial of Phase 3, while most familiarized *Ss* (Groups E-F and C-F) had acquired this skill by the beginning of Phase 3. Although any recognizable pronunciation was accepted by *E* as a correct response (provided *S* was consistent in pronunciation), non-familiarized *Ss* made fewer fragmentary 'guessing' responses. A similar experiment has been done in serial learning by Cieutat.[18] The associations his *Ss* learned to an item were selected from those actually given by other *Ss*, thus 'belonging' in the sense discussed above. His techniques in training and familiarization were different, but the *n*-factor was about as effective as the *m*-factor. Although the present results are consistent with other studies in this area, there are many questions remaining to be answered about the roles played by meaningfulness and familiarity in verbal behavior. Schulz and Tucker have reported evidence that the facilitation of the S-term, found in the experiments of Gannon and Noble and Hakes, is dependent on the method of two-term pronunciation and interacts with length of the anticipatory interval,[19] so instructions and temporal factors may be relevant to R-term effects as well. Finally, the fact that a recent study by Noble revealed both S-term and R-term facilitation in paired-associate CVC material points to the need for further attempts to control the *n*-factor in experimenting on induced meaningfulness.[20] Evidence that Groups C-F and E-F reacted so differently to familiarization leads us to reserve judgment on the validity of the hypothesis of Cieutat, Stockwell and Noble and of Underwood and Schulz.[21]

---

[17] Noble, *op. cit.*, 1963, Ch. 4.

[18] V. J. Cieutat, Induced meaningfulness in serial verbal learning, *Percept. Mot. Skills,* 14, 1962, 373-374.

SUMMARY

Two experiments tested the hypothesis that number of associations is the major factor producing the facilitative influence of meaningfulness (*m*). In Experiment I, 20 *S*s learned 0, 3, 6, or 9 associations to low-*m* items which later appeared in the response-positions of a paired-associate task. In Experiment II, testing for the influence of differential familiarization (*n*) on the response term, an Experimental Group ($N = 48$) replicated Experiment I, an Experimental-Familiarized Group ($N = 24$) received in addition 32 exposures (*n*) to the untrained response-terms, a Control Group ($N = 48$) had normative-*m* terms substituted for items with built-in *m*, and a Control-Familiarized Group ($N = 24$) was treated like the Controls except for 32 exposures to the low-*m* items.

Both experiments supported the hypothesis that added associations can produce the *m*-effect, but Experiment II revealed a significant interaction between *m* and *n*. Performance on the Experimental-Familiarized group was independent of *m*, all others increasing significantly. Familiarized groups were generally more proficient at low levels of *m*, whereas control groups were superior at higher levels. Complexities in the data preclude definite identification of the role of *n*.

---

[19] Schulz and Tucker, *op. cit.*

[20] Noble, *op. cit.*, 1963, Ch. 4.

[21] Cieutat, Stockwell, and Noble, *op. cit.*, 201; Underwood and Schulz, *op. cit.*, 126-127.

FACTOR-STRUCTURE OF SEMANTIC DIFFERENTIAL RESPONSES TO VISUAL FORMS AND PREDICTION OF FACTOR-SCORES FROM STRUCTURAL CHARACTERISTICS OF THE STIMULUS-SHAPES

By Lois L. Elliott, Brooks Air Force Base, Texas, and
Percy H. Tannenbaum, University of Wisconsin

The work presented in this paper combines two previously unrelated approaches to psychological phenomena. It utilizes stimulus-materials and certain theoretical concepts from the psychophysics of visual form-perception as well as techniques and hypotheses concerning meaning and language which are pertinent to the semantic differential.

Previous investigators, hypothesizing that perceptual responses are determined by physical characteristics of the stimulus which are theoretically amenable to exact specification, have attempted to develop a psychophysics of form-perception.[1] For example, Attneave demonstrated that a large proportion (90%) of the variance of scaled judgments of stimulus-complexity could be related to certain physical characteristics by multiple correlation techniques.[2] Arnoult, in turn, showed that not only 'complexity,' but also judgments of size, familiarity, and meaningfulness could be related to physical characteristics of the stimuli although not always with as much success.[3] Recently, Edelman, Karas, and Cohen, in exploring judged complexity and latency of free association for randomly generated symmetrical and asymmetrical shapes, obtained higher association-latencies for more complex shapes and lower latencies for symmetrical shapes (when total number of sides was equated).[4]

The other aspect of this study, the semantic differential, which was originated at

---

* Received for publication May 4, 1962. This research was initiated in 1957 under Project 7706 at the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. A modified version having the same title has been published as School of Aerospace Medicine Technical Report 62-8, November 1961. The Institute of Communications Research, University of Illinois also supported the research; the authors wish to thank Edward Ware of the University of Illinois for his aid with ILLIAC computations. Appreciation is expressed to Dr. Harry Hughes of the School of Aerospace Medicine for providing computer time and other assistance in the later states of the analysis. Dr. Elliott is now at Central Institute for the Deaf, St. Louis, Missouri.

[1] Among others, see J. J. Gibson, *The Perception of the Visual World*, 1950, 59-76.

[2] Fred Attneave, Physical determinants of the judged complexity of shapes, *J. exp. Psychol.*, 53, 1957, 221-227.

[3] M. D. Arnoult, Prediction of perceptual responses from structural characteristics of the stimulus, *Percept. mot. Skills*, 11, 1960, 261-268.

[4] S. K. Edelman, G. G. Karas, and B. J. Cohen, The relative contributions of complexity and symmetry to the perception of form, Paper presented at Midwestern Psychological Association, May 1961.

the University of Illinois by Osgood, has been directed at determining the underlying dimensions of meaning.[5] Although experimental results have varied, particularly as new concepts and new cultures have been investigated, the dimensions most frequently recurring have been those of evaluation, potency, and activity.[6]

Hochberg, using a modified version of the semantic differential to evaluate responses to the words *takete* and *uloomu* as well as to the graphic figures which are usually matched with these words, obtained evidence of greater similarity in semantic differential response to matching pairs of word and figure.[7]

Research in audition has investigated semantic dimensions of complex sounds and their associated physical correlates. Solomon applied this approach using both Sonar signals and systematically varied sounds. His results showed that semantic judgments were most strongly related to frequency spectra although rhythm of amplitude modulations was also important.[8]

The present study was designed first, to locate the major semantic dimensions applicable to a population of randomly generated, nonsense-shapes. The second purpose was to determine the relationship of these dimensions or factors to physical characteristics of the stimuli by multiple correlation techniques.

Method: (1) Stimuli. The shapes used in this study (70 in number) were selected from the 72 shapes employed by Arnoult and Elliott[9] Thirty-six of the shapes were constructed by Method 1 as outlined by Attneave and Arnoult, while the remaining 34 were constructed by Method 2.[10] Since Arnoult has recently demonstrated that the shapes constructed by Method 1 and those prepared by Method 2 are responded to by *Ss* as partially different sets,[11] the stimuli in the present study were divided for analytical purposes into two groups and designated Sample 1 and Sample 2.[12]

Photographs of the shapes, measuring 3½ × 5 in., were prepared and assembled in random order in individual test-booklets equipped with hard covers. Since there was inadequate time for each *S* to judge all shapes, the total stimulus-series was divided into three parts: (a) three groups of 20 shapes with physical parameters distributed across the three groups; (b) a pair of two very dissimilar shapes (practice shapes); and (c) a residual set of 10 shapes. Each *S* judged 32 shapes: (a) the pair of practice shapes for which judgments were discarded; (b) 20 shapes from 1 of the 3 groups; and (c) the 10 shapes of the residual group.

---

[5] C. E. Osgood, The nature and measurement of meaning, *Psychol. Bull.*, 49, 1952, 197-237.

[6] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, *The Measurement of Meaning*, 1957, 1-75; esp. 72f.

[7] J. E. Hochberg, Psychophysics and stereotype in social perception, in Muzafer Sherif and M. O. Wilson, (eds.), *Emerging Problems in Social Psychology*, 1957, 134.

[8] L. N. Solomon, Semantic approach to the perception of complex sounds, *J. acoust. Soc. Amer.*, 30, 1958, 421-425; Search for physical correlates to psychological dimensions of sounds, *ibid.*, 31, 1959, 492-497; Semantic reactions to systematically varied sounds, *ibid.*, 31, 1959, 986-990; In another modality Tannenbaum has used a similar approach in the study of color stimuli (unpublished research).

[9] Arnoult, *op. cit.*, 262; L. L. Elliott. Reliability of judgments of figural complexity. *J. exp. Psychol.*, 56, 1958, 335-338.

[10] Fred Attneave and M. D. Arnoult, The quantitative study of shape and pattern perception, *Psychol. Bull.*, 53, 1956, 452-471.

[11] Arnoult, *op. cit.* 264.

[12] Sample 1 includes Symmetry and Curvature as variables, while Sample 2 does not (see footnote 11).

Measurable physical characteristics of the shapes included the following variables: Perimeter (P), Area (A), Perimeter squared divided by Area (P^2/A), Angular Variability (AV), Total Sides (TS), Independent Sides (IS), Symmetry (Sy) and Curvature (C).[13] These measures were defined in the same manner reported by Attneave.[14]

(2) Subjects. Participating in the experiment were 300 basic trainees from the Air Force, screened to possess aptitudes roughly equivalent to those of a college population. Previous work had shown that ratings of shape-complexity made by basic airmen are highly consistent with paired-comparisons made by college graduates.[15] The *Ss* were tested in groups of 15. One hundred *Ss* judged each subset of 20 stimuli, while all 300 *Ss* judged the remaining 10 shapes.

(3) Scales. Twenty bi-polar adjective-scales representing factors obtained in other research with the semantic differential were selected as being appropriate to the stimulus-series under study. A seven-point rating-form was used.

(4) Procedure. Booklets were prepared containing separate pages for judgments of each shape. Instructions introduced the experiment as a research designed to improve terminology with which Air Force navigators might discuss radar displays. The *Ss* were required to examine each photograph in their booklet before starting to rate the practice-shapes.

Results. Analysis of results consisted of two phases: obtaining the factor structure of the semantic judgments and determining the relationship of factor scores to the physical measurements.

The 20 × 20 intercorrelation based on standardized mean rating-scores for each of the shapes (each correlation coefficient based on 70 pairs of standardized mean ratings) was factored by the centroid method and rotated by the quartimax technique.[16] This was repeated three times: independently for the shapes in Samples 1 and 2 and finally for the total sample. Extraction of factors in each sample was halted at four factors, since approximately all of the estimated common variance had been extracted. Results of the three analyses are shown in Table I.

Since, with few exceptions, the three resulting sets of factors are in excellent agreement, they will be discussed jointly.

The first, Factor I, represents a complexity-activity dimension with *fancy-plain*,

---

[13] Total Sides refers to the number of lines joining separate points in the original plotting of the figure. The measure of Independent Sides differs from this (a) when the shape is symmetric or (b) when, for example, three plotted points fall on a straight line so that only one line is required to connect them instead of two. Angular variability refers to the mean of the difference between size of external angles of the shape measured consecutively. For example, a regular figure such as a star would have 0.0 Angular Variability, while a highly irregular figure would have a considerable larger AV score. Area and Perimeter are self-explanatory. The measures of Symmetry and Curvature apply only to Sample 1 and are dichotomous measures.

[14] Attneave, *op. cit.*, 222-223

[15] Elliott, *op. cit.*, 336.

[16] Factoring was initiated prior to general availability of the varimax technique.

TABLE I

Quartimax Rotations of Factor Matrices for Sample 1, Sample 2, and the Total Sample

Scale	Factor loadings												Communalities		
	Factor I			Factor II			Factor III			Factor IV					
	Sample 1	Sample 2	Total	Sample 1	Sample 2	Total	Sample 1	Sample 2	Total	Sample 1	Sample 2	Total	Sample 1	Sample 2	Total
Pleasant —unpleasant	19	00	41	−94	−90	−86	07	13	17	−19	−07	14	97	84	97
hard —soft	−42	−01	−61	17	67	33	−34	−38	−42	79	41	−55	96	78	98
active —passive	−94	−95	−97	02	−08	00	21	14	15	16	03	−12	97	95	99
careful —careless	21	31	45	−90	−90	−83	−09	−02	−03	15	02	−14	90	92	93
warm —cool	59	09	67	−41	−30	−40	−02	16	16	−61	−54	47	90	43	88
usual —unusual	79	88	89	−48	−17	−31	11	26	12	10	−02	−11	89	89	92
ugly —beautiful	09	58	12	95	75	89	−06	−16	−19	09	−01	−08	93	94	86
light —heavy	−17	−60	−37	−08	−24	−15	93	76	88	−08	00	−03	91	99	95
slow —fast	79	89	90	16	10	09	−39	−36	−30	−29	−06	21	91	96	97
sober —drunk	47	61	61	−74	−52	−59	12	00	02	30	16	−35	89	69	86
fancy —plain	−97	−98	−98	03	−01	−02	−14	−06	−04	04	−07	08	97	99	99
large —small	−07	41	00	02	28	14	−93	−79	−90	10	−02	00	88	89	84
good —bad	16	00	38	−94	−88	−88	04	08	12	−15	02	11	96	80	95
masculine—feminine	−14	57	−08	11	42	30	−61	−61	−74	63	15	−45	82	91	86
excitable —calm	−96	−98	−98	20	07	11	07	06	08	07	12	00	98	99	99
stable —unstable	73	90	86	−51	−09	−30	−30	−29	−32	25	01	−16	97	92	97
colorful —colorless	−93	−98	−94	−21	−07	−19	−15	03	01	01	−05	11	94	99	95
horizontal—vertical	58	39	59	−04	−23	−10	−24	−11	−10	−34	−09	23	52	23	43
simple —complex	89	92	95	−27	−09	−14	23	26	18	−12	07	−03	95	94	97
rounded —angular	49	70	70	10	−18	−09	03	−05	18	−79	−60	62	86	91	92
% of variance accounted for	51	56	61	23	32	24	19	7	11	7	5	4			

(Decimal points have been omitted.)

excitable-calm, active-passive, simple-complex, and *colorful-colorless* having highest loadings. Scales having substantial loadings include *slow-fast, usual-unusual, stable-unstable, horizontal-vertical, sober-drunk,* and *rounded-angular. Warm-cool* has moderate loadings for Sample 1 and for the total group but fails to load for Sample 2. In contrast *heavy-light* has a moderate loading in Sample 2 but fails to contribute appreciably in either Sample 1 or the total group.

Factor II represents an esthetic-evaluative dimension with *ugly-beautiful, good-bad, pleasant-unpleasant,* and *careful-careless* having highest loadings. In addition, *hard-soft* has a moderate loading for Sample 2.

Factor III refers to size or potency with *light-heavy* and *large-small* having high loadings. In addition, *masculine-feminine* has moderately high loadings in all three analyses. Considering the association of *light, small,* and *feminine,* this might also be interpreted as a dimension of femininity.

Factor IV, less clearly defined, denotes a 'hardness-angularity' dimension. Defining scales were *hard-soft, rounded-angular, warm-cool,* and for Sample 1, *masculine-feminine.* Another interpretation might contrast this with Factor III and call it a masculinity factor with *hardness, angularity,* and *coolness* mutually associated. Supporting this suggestion is the fact that four of the six scales involved have high loadings for one factor and essentially zero loadings for the other. Intercorrelations of Factors III and IV are −0.16 for Sample 1 and −0.28 for Sample 2 which are not significantly different from zero.

These factors agree fairly well with those found in semantic differential analyses of other materials. Factor I is roughly an 'activity' factor; Factor II an 'evaluation' dimension, while Factor III and perhaps IV could be considered as 'potency' factors.

In exploring the relationship between the factors and the physical measures, it was first necessary to obtain factor-scores for the stimuli. This was accomplished separately for the two samples by regressing semantic differential scores (standardized within each subsample) on factor-loadings according to the formula $P = S^1F(F^1F)^{-1}$ where 'S' contains mean rating-scores standardized over 70 shapes, 'F' contains rotated factor loadings, and 'P' contains estimated factor scores. (Shapes with high and low factor scores are shown in Figs. 1 and 2.) Regression coefficients and multiple correlations were computed within each sample by regressing factor scores on the physical measurements. (Since both scores and physical measurements were standardized within each group, the regression coefficients are actually beta coefficients.) Results are shown in Tables II and III.

Multiple correlations for three factors of Sample 1 are higher than the corresponding ones for Sample 2. This result could be due to the different methods of constructing the shapes, to the presence of symmetry and curvature variables in Sample 1 but not in Sample 2, or to both possibilities. Multiple correlation coefficients for Factor I (complexity-activity) are roughly equivalent and reasonably substantial in both groups. Factor III (size-potency) also has fairly good prediction from the physical variables, although the multiple correlation is higher in Sample 1. Examination of Tables II and III shows that for the two factors (Factors II and IV), which

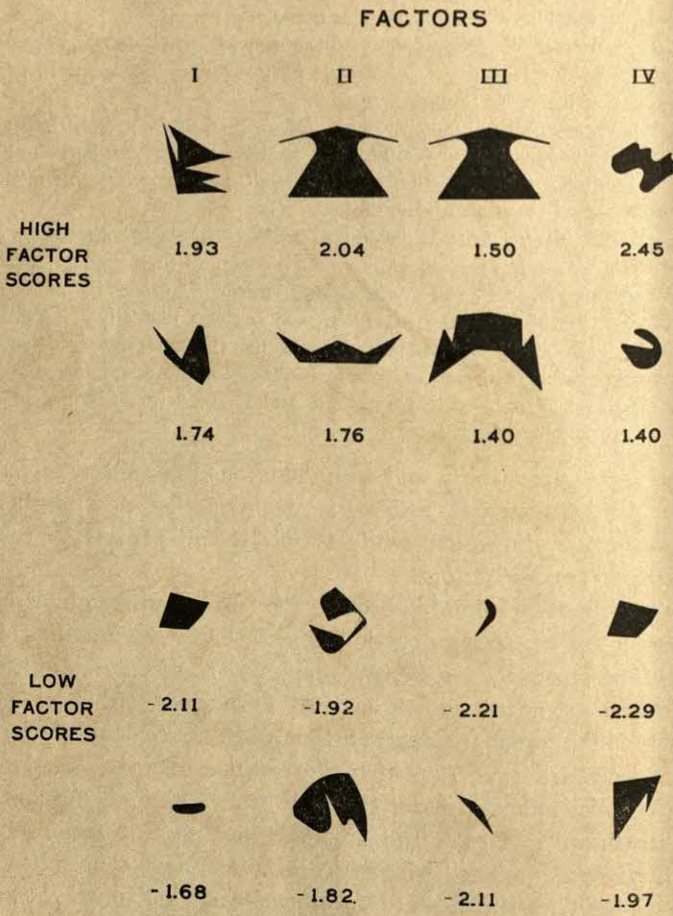


FIG. 1. REPRESENTATIVE SHAPES WITH HIGH AND LOW FACTOR SCORES: SAMPLE 1.

have very low multiple correlations in Sample 2 as compared to Sample 1, the variables of symmetry and curvature have relatively large beta coefficients in Sample 1. Results in Table III indicate that the relative magnitudes of beta coefficients for all factors differ considerably between Samples 1 and 2. Presumably this also is attributable to the symmetry and curvature variables in Sample 1. Other investigators have found that shapes possessing symmetry constitute a distinct group.[17]

Comparisons of these multiple correlations with those obtained by Arnoult are shown in Table IV.[18] Factor I of this study corresponds to

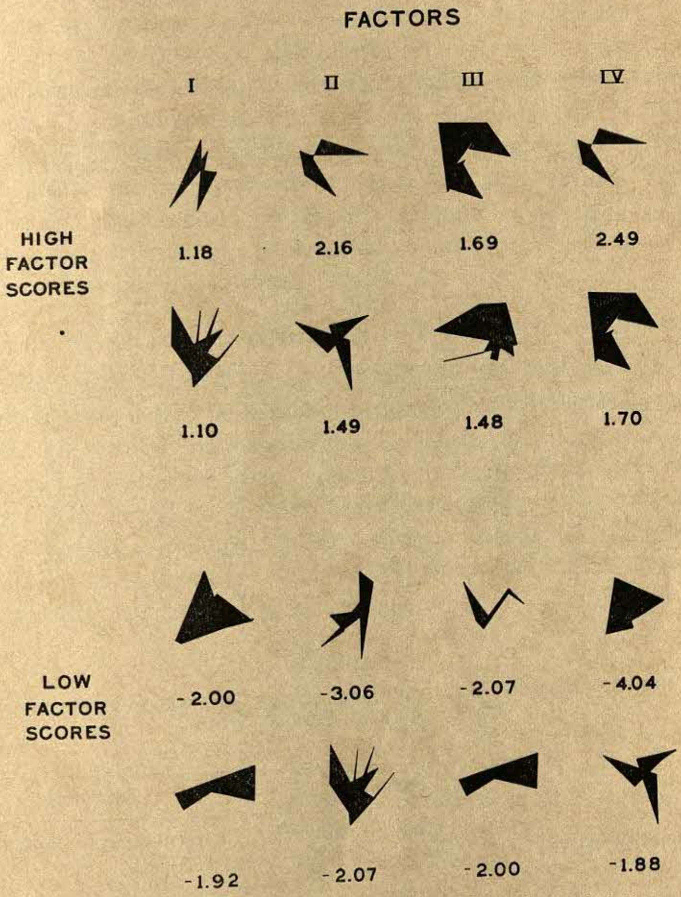


FIG. 2. REPRESENTATIVE SHAPES WITH HIGH AND LOW FACTOR SCORES: SAMPLE 2.

Arnoult's complexity judgments, while Factor III is related to his judgments of size. There are two major differences between the two studies. In Arnoult's work, the multiple correlations involve a single perceptual judgment while in the present study the factor scores are obtained by compounding judgments on 20 adjectival scales. Secondly, Arnoult selected the best predictor variables while in the present study all available physical measurements entered the analyses. In view of these differences it is remarkable that nearly identical multiple correlations were obtained. While

---

[17] Edelman, Karas, and Cohen, *op. cit.*

[18] Arnoult, *op. cit.*, 265.

TABLE II

MULTIPLE CORRELATIONS FOR EACH FACTOR REGRESSED ON PHYSICAL MEASUREMENTS

	R_{mult}	
	Sample I	Sample II
Factor I	.87	.89
Factor II	.84	.35
Factor III	.93	.83
Factor IV	.95	.42

TABLE III

BETA COEFFICIENTS FOR EACH FACTOR REGRESSED ON PHYSICAL MEASUREMENTS

	Factor I		Factor II		Factor III		Factor IV	
	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
P	−1.01	.24	−1.32	.45	.87	.07	−1.09	.54
A	.31	−.31	.51	−.67	−.05	.49	.21	−.22
P^2/A	.77	.18	.59	−.59	−.56	−.37	.53	−.33
AV	.49	.30	.43	.12	.04	.26	.34	.14
TS	−.32	.28	−.09	−.12	.00	.03	−.17	−.26
IS	1.08	.04	.02	−.04	.21	.41	.83	.29
Sy	.70		.97		.30		.87	
C	−.20		−.20		−.01		.82	

TABLE IV

COMPARISON BETWEEN MULTIPLE CORRELATIONS OBTAINED BY ARNOULT AND OBTAINED IN THE PRESENT STUDY

	R_{mult}			
	Sample I		Sample II	
	Present study	Arnoult	Present study	Arnoult
Complexity-Activity	.87	.93	.89	.91
Size	.93	.96	.83	.86

Arnoult's coefficients are slightly larger than ours, he, too, found poorer prediction for Sample 2.

Discussion. Factor analysis in this study appears to have two major implications. The factors resulting from programmed rotational procedures correspond reasonably well to those obtained in semantic differential analyses of other materials. This contributes additional support to the thesis that the same meaningful dimensions underly a large portion of human behavior. The second finding, that of very similar factor structures in separate analyses for the two stimulus-samples, provides strong evidence that the results obtained are not chance occurrences. When viewed in the context of the multiple correlations, the results imply that the semantic dimensions may be generalized to other populations of visual stimuli more readily than can the predictive equations based on physical measurements.

The findings of the multiple correlations suggest the possibility that in judgments of nonsense-shapes the portion of semantic space enveloped by the set of complexity judgments (Arnoult's) roughly coincides with that denoted by the complexity-activity factor—and similarly for the size-judgments and the factor of size-potency. It should be stressed that this hypothesis would pertain to nonsense-shapes and would not necessarily apply to more meaningful stimuli.

SUMMARY

Two samples of nonsense-shapes, containing 36 and 34 stimulus-shapes, respectively, were judged on 20 semantic differential scales. Separate factor analyses and rotations produced highly similar factor structures for the two samples and the total group. Four factors were obtained, representing dimensions of complexity-activity, esthetic-evaluation, size-potency and hardness-angularity. Multiple regression solutions, predicting factor-scores from physical measurements, indicated better prediction for the sample in which symmetry and curvature were varied.

THE EFFECT OF BODY-POSITION ON MERIDIONAL VARIATIONS IN SCOTOPIC ACUITY

By S. M. LURIA, Groton, Connecticut

It is well established that visual acuity is better for vertical or horizontal (*HV*) test-objects than it is for those which are diagonally (*D*) oriented,[1] and that the variability of judgments associated with *D*-lines is greater than that with *HV*-lines.[2] Nevertheless, the basis of this phenomenon has remained obscure.

Originally, psychological explanations were advanced. In 1893, when Jastrow found that the *O*s made more precise judgments of *HV*-lines than of *D*-lines, he attributed it to the greater familiarity of *HV*-objects.[3] Later, Koffka wrote that we develop a visual framework for space.[4]

Recently, physiological approaches have claimed more attention. Brown has suggested that eye-movements are responsible for the effect,[5] but Higgins and Stultz have found it even with exposure-times of less than 1 m.sec., thus eliminating eye-movements as the cause.[6] Leibowitz reported acuity-thresholds for various combinations of pupil-diameter and retinal illumination.[7] His data showed systematic changes as the diameter of the pupil varied which, as he thought, pointed to a dioptric origin of the differences. The persistence of the phenomenon with large pupil-diameters and high illuminance also demonstrated the importance of retinal or post-retinal factors. Ogilvie and Taylor supported Leibowitz's contention that dioptric factors could not provide a full explanation, since they obtained the effect for *O*s with no astigmatism.[8]

---

* Received for publication May 19, 1962. From the U.S. Naval Medical Research Laboratory.

[1] G. C. Higgins and K. Stultz, Visual acuity as measured with various orientations of a parallel-line test object, *J. opt. Soc. Amer.*, 38, 1948, 756-758; Variations of visual acuity with various test-object orientations and viewing conditions, *ibid.*, 40, 1950, 135-137; Herschel Leibowitz, Some observations and theory on the variation of visual acuity with the orientation of the test object, *ibid.*, 43, 1953, 902-905; J. C. Ogilvie and M. M. Taylor, Effect of orientation on the visibility of fine wires, *ibid.*, 48, 1958, 628-629.

[2] A. M. Rochlin, The effect of tilt on the visual perception of parallelness, this JOURNAL, 68, 1955, 223-236; Leibowitz, Some factors influencing the variability of vernier adjustments, this JOURNAL, 68, 1955, 266-273; E. A. Craig and M. Lichtenstein, Visibility-invisibility cycles as a function of stimulus-orientation, this JOURNAL, 66, 1953, 554-563; J. W. Onley and John Volkmann, The visual perception of perpendicularity, this JOURNAL, 71, 1958, 504-516.

[3] Joseph Jastrow, On the judgment of angles and positions of lines, this JOURNAL, 5, 1893, 214-248.

[4] Kurt Koffka, *Gestalt Psychology*, 1935, 190-210.

[5] W. R. J. Brown, Eye movements and the variation of visual acuity with test object orientation, *J. opt. Soc. Amer.*, 39, 1949, 1057.

[6] Higgins and Stultz, *op. cit.*, *J. opt. Soc. Amer.*, 40, 1950, 135-137.

[7] Leibowitz, *op. cit.*, *J. opt. Soc. Amer.*, 43, 1953, 902-905.

[8] Ogilvie and Taylor, *op. cit.*, 628-629.

Weymouth has contended that the phenomenon is optical,[9] and has called attention to older demonstrations that point-sources are not focused as points but as stars with prominent horizontal and vertical rays.[10] This fact, as he believed, completely accounted for the effect. The investigation of other possibilities has, however, continued. Nachmias has studied the effect of directional non-uniformities of eye-movements, concluding that these did not account for the phenomenon in his *O*s.[11] Most recently, Foley has found that the phenomenon cannot be attributed to differential retinal summation.[12]

Although most investigators prefer other than psychological explanations, Hebb has proposed a comprehensive theory of the development of behavior which would account for the phenomenon in a way not incompatible with earlier ideas.[13] Discussing his experiment in this area, Rochlin wrote:

> The greatest amount of perceptual experience with spatial relations occurs in life-situations with structures in the horizontal and vertical positions. Behavior is, in general, oriented in relation to the main axes of space, and sensitivity of perception may, therefore, be expected to be greater for objects located in line with these axes than for objects in oblique positions.[14]

Onley and Volkmann concluded from their study that:

> The subjective axes of horizontal and vertical play a conspicuous role in the adjustment of two lines to be mutually perpendicular. . . . All *S*s apparently used the primary reference axes of the subjective horizontal and vertical in the judgment of the angular orientation of lines.[15]

Since there is still a tendency in some quarters to think in psychological terms and in view of the disappointing results of the ocularly oriented studies. the present experiment was undertaken to determine the effects of the position of *O*'s head and body on acuity for *HV*- and *D*-targets. Ogilvie and Taylor have stated that when the head is tilted 45° from the vertical, diagonal wires are more easily visible than horizontal ones,[16] but the present study aims to present more detailed information than they gave, and, in addition, to attempt to reduce the visual cues for verticality. If such psychological factors are important, we should expect, for example, the differences between *HV*- and *D*-acuities to decrease when there are conflicting cues as to the position of the vertical. Such changes would indicate that the phenomenon is not wholly a physiological one.

---

[9] F. W. Weymouth, Stimulus orientation and threshold: An optical analysis, *Amer. J. Ophthal.*, 48, 1959, Pt. II, 6-10.

[10] J. P. C. Southall, *Helmholtz's Treatise on Physiological Optics*, I, 1924, 416-443.

[11] Jack Nachmias, Meridional variations in visual acuity and eye movements during fixation, *J. opt. Soc. Amer.*, 50, 1960, 569-571.

[12] P. J. Foley, Stimulus orientation and retinal summation, *J. opt. Soc. Amer.*, 52, 1962, 474-475.

[13] D. O. Hebb, *The Organization of Behavior*, 1949, 91-95.

[14] Rochlin, *op. cit.*, 235.

[15] Onley and Volkman, *op. cit.*, 515.

[16] Ogilvie and Taylor, Effect of length of line on the visibility of a fine line, *J. opt. Soc. Amer.*, 49, 1959, 898-901.

Method. Acuity was measured with a graded set of high contrast checkerboard-targets using the method of constant stimuli. Each target was presented in random order 10 times per session; there were three sessions for each condition. The targets were inserted in a circular frame with a diameter of 10° of visual angle, centered 10° to the left of a small red fixation-cross. The frame was 10 ft. from *O* and illuminated to 5 log μμL. Black curtains were hung around the room and spread on the floor in such a manner that only the fixation-light, the frame, and a wooden rod under the frame were visible during the exposure-periods of 3 sec. The rod under the frame was a simulated frame-support and was intended to provide a cue to the vertical. Between exposures, a circular white shutter blocked the field except for the fixation-light. Since, the shutter was matched in brightness to the unresolved targets, *O*'s level of adaptation was not changed by its rise and fall.

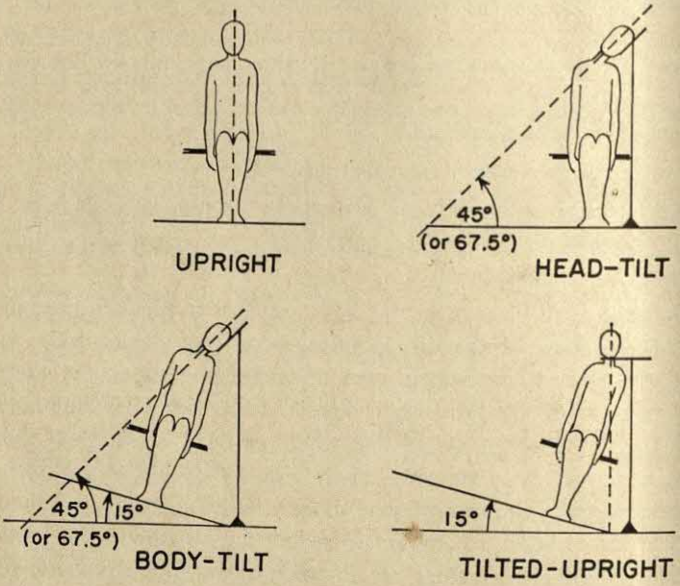


FIG. 1. POSITIONS OF THE *O*S

There were three conditions: (1) the upright position in which *O* sat in a chair in the usual way; (2) the head-tilt position in which *O* sat in the same chair, but with his head resting on a tilted support; and (3) the body-tilt position in which the chair was placed on a platform tilted at 15° to the floor and with *O*'s head again resting on the tilted head-rest (see Fig. 1).

In Conditions 1 and 2, the white rod below the target-frame was perpendicular to the floor; in the third condition, however, the rod was attached to the frame at an angle to the floor. Two angles of tilt from the vertical were investigated in separate experiments (45° and 22.5°), and these two degrees of tilt were used for the target 'support.' When the head was tilted, the fixation-light was set at about the 1:30 o'clock position in respect to the target, while in the upright condition, it was set at the 3:00 o'clock position.

The chair and head-rest were padded with 3 in. of foam rubber. It has been reported that such padding significantly reduces the kinesthetic sensations which give cues to the position of the true vertical.[17] Since there is evidence that vision is the dominant factor in the orientation of normals,[18] it was expected that the padding, combined with the rod under the target-frame, would permit some manipulation of *O*'s perception of verticality.

Before the observations began, *O* was dark-adapted for 15 min. and then adapted to the target-luminance for 4 min. while sitting in the test-position. This procedure presumably served also to adapt the *O*, to the kinesthetic sensations to some extent, for it has been reported that there is adaptation to bodytilting, at least during a brief period.[19] The observations were made with the right eye.

TABLE I

RESULTS OF EXAMINATIONS FOR ASTIGMATISM WITH *O*s WEARING THEIR GLASSES

O	Diopters	Axis
JK	0.00	—
AR	+0.25	180
MS	+0.50	100
JV	+0.00	—

Observers. Four *O*s were tested, two men (*MS* and *JV*) and two women (*JK* and *AR*), ranging in age from 28–55 yr. All except *JV* were experienced *O*s. Two were emmetropic (*MS* and *AR*), one was hyperopic (*JK*), and one was myopic (*JV*). The ametropes wore their glasses during the experiment. All were examined for astigmatism with their corrections, with the results shown in Table I. It would be expected that *AR* might have slightly better acuity for horizontal lines (although 0.25 is considered negligible), while *MS* would have slightly better acuity for lines 10° clockwise from vertical.

Results. The various thresholds of scotopic acuity are presented in Table II in terms of the size in millimeters of the squares on the threshold-target.[20] In the upright position, all *O*s had better acuity for the *HV*- than for the *D*-orientation. The differences between the two thresholds are highly significant ($p = 0.01$)[21] and also sizable for all *O*s except *AR*.

---

[17] C. W. Mann and H. J. Dauterive, Jr., The perception of the postural vertical: The modification of non-labyrinthine cues, U.S.N. School of Aviat. Med. and Res., Nav. Air Sta., Pensacola, Jan. 25, 1950.

[18] M. E. Bitterman and Philip Worchel, The phenomenal vertical and horizontal in blind and sighted subjects, this JOURNAL, 66, 1953, 598-602.

[19] C. W. Mann, G. E. Passey, and R. K. Ambler, The perception of the vertical: VII. Effect of varying intervals of delay in a tilted position upon the perception of the postural vertical, U.S.N. School of Aviat. Med. and Res., Nav. Air Sta., Pensacola, Jan. 25, 1950.

[20] These numbers have not been converted to the more usual indices of acuity because the target-sizes are much more easily grasped.

[21] The significance of the differences between the limens was computed using the correction-factor developed by H. von Schelling (S. W. Smith, A. Morris, and F. L. Dimmick, Effects of exposure to various red lights upon subsequent dark adaptation measured by the method of constant stimuli, *J. opt. Soc. Amer.*, 45, 1955, 502-506).

With *O*'s head tilted 45°, acuity was always better for the *D*- than for the *HV*-orientation, but the average difference between thresholds is less than half the difference in the upright position (only *AR*'s difference is greater than in the first condition), and only two of them are significant. Furthermore the thresholds are higher than in the upright condition; in every case, the better threshold in the head-tilt condition is worse than the better threshold in the upright condition.

With both head and body tilted, the thresholds were again lower for the *D*- than for the *HV*-orientation for three *O*s, but there was a reversal for *JV*. His difference, however, is the only one not significant. These differences, furthermore, are in every case larger than in the head-tilt condition. Thus, in the upright condition, the average difference between means was 2.0 mm.; in the head-tilt condition, it dropped to 0.9 mm., and in the body-tilt condition it rose to 1.22 mm. The difference in the upright condition is significantly greater ($p = 0.05$) than either of the other differences, but the body-tilt difference is not significantly greater than the head-tilt difference.

TABLE II

ACUITY-THRESHOLDS EXPRESSED IN MM. OF TARGET-SIZE

O		Upright			Head-tilt			Body-tilt		
		HV		*D*	*HV*		*D*	*HV*		*D*
JK	T	10.1		13.1	12.7		11.4	12.5		10.7
	σ	0.8		0.7	1.3		1.2	0.7		0.9
	D		3.0*			1.3*			1.8*	
AR	T	10.9		11.7	13.3		11.6	13.7		11.8
	σ	0.9		0.6	0.9		1.5	0.7		2.9
	D		0.8*			1.7*			1.9*	
MS	T	14.0		15.5	15.7		15.5	15.2		14.5
	σ	0.9		1.0	1.2		1.9	1.4		1.2
	D		1.5*			0.2			0.7†	
IV	T	16.9		19.6	18.3		17.9	17.9		18.4
	σ	1.5		2.3	1.4		1.4	1.9		1.5
	D		2.7*			0.4			0.5	

* Significant at the 1% level.
† Significant at 5% level.

After these data had been collected, acuity-thresholds obtained with the head tilted only 22.5° from the vertical were compared for the head-tilt and body-tilt positions. These results are shown in Table III. For three *O*s, acuity was better for *HV* in both conditions, but *AR* had better acuity for the *D*-orientation.[22] There was no uniform change in magnitude of difference between means; two decreased in the body-tilt condition and two increased. Except for *JV*, the thresholds were again worse than in the upright position.

---

[22] This fact must be related to her small difference between thresholds in the upright condition. One would suppose it to be an indication of an astigmatic condition in a diagonal axis, but the ophthalmic examination revealed no such finding.

Discussion. It is obvious that postural variations as well as target-orientation affected acuity. According to the 'sensory-tonic' theory,[23] it would be expected that if motor activity increased, perceptual activity would decrease, and the introduction of unusual postures has, indeed, reduced acuity in all cases in the first part of the study and in all cases but one (*JV*) in the second part. It should be noted, however, that *JV* was the most inexperienced *O*, and his improved threshold in the second part of the study may reflect the benefit of the practice in the first part. Of more interest to us is the fact that acuity for the two target-orientations was differentially affected by changes in posture. Although acuity was (with one exception) better for targets parallel to the median plane of *O*'s head, the difference between the thresholds for targets parallel to this plane and diagonal to it varied with posture. The largest average difference occurred with head and body upright, the smallest difference with only the head tilted; an intermediate difference resulted with both head and body tilted. Apparently, when head and body are in line with the apparent vertical, the orientational phenomenon occurs with greater effect than when these factors are at odds.

TABLE III

VISUAL ACUITY THRESHOLDS EXPRESSED IN MM. OF TARGET-SIZE

O		22.5° head-tilt			Body-tilt*		
		HV		*D*	*HV*		*D*
	T	10.6		11.9	10.3		11.4
JK	σ	0.8		0.7	1.0		0.7
	D		1.3†			1.1†	
	T	12.5		11.6	12.4		11.7
AR	σ	0.7		1.1	1.0		1.0
	D		1.1†			0.7‡	
	T	15.2		16.7	14.8		16.8
MS	σ	1.3		1.6	0.8		1.3
	D		1.5†			2.0†	
	T	15.8		17.7	16.1		18.4
JV	σ	1.5		1.3	2.0		2.3
	D		1.9†			2.3†	

* Body-tilt 15° from horizontal; head-tilt 22.5° from vertical.
† Significant at 1% level.
‡ Significant at 5% level.

---

[23] Heinz Werner and Seymour Wapner, Toward a general theory of perception, *Psychol. Rev.*, 59, 1952, 324-338.

It may be asked whether these results are due to a partial success in altering *O*'s perception of the vertical, or whether they are a function of the amount of twisting of the neck or other discomforts. One would suppose that the discomfort of sitting in the tilted chair would be at least as great as resting the head at a 45° angle, and therefore that the differences between means should have been the same. Similarly, it would seem that tilting the head a mere 22.5° would be only negligibly uncomfortable; yet the differences between the thresholds are not uniformly greater in the 22.5° head-tilt condition than in the corresponding body-tilt condition, which was somewhat uncomfortable.

TABLE IV

VISUAL ACUITY EXPRESSED IN MM. OF TARGET-SIZE

O		Upright			Tilted upright		
		HV		*D*	*HV*		*D*
JK	T	9.8		12.4	10.4		13.2
	σ	0.6		0.5	0.6		1.1
	D		2.6			2.8	
AR	T	10.8		11.4	11.2		11.7
	σ	0.8		0.7	0.7		1.2
	D		0.6			0.5	
JV	T	16.8		17.8	16.2		18.0
	σ	2.3		2.2	1.7		1.1
	D		1.0			1.8	
SL	T	16.6		20.1	17.7		21.2
	σ	2.0		3.0	2.2		2.1
	D		3.5			3.5	

As a check, however, thresholds in the upright condition were compared with those obtained with four *O*s sitting on the tilted chair but with their heads held in the upright position by a chin-rest and head-frame (tilted-upright position).[24] These results are given in Table IV. For three *O*s, there was virtually no change in the difference between thresholds. For *JV*, this difference increased from 1.0 to 1.8 in the tilted-upright position, but it will be noted that this change is in the opposite direction from those found in the main experiment; furthermore, the difference of 1.0 is, for some reason, much smaller than his original difference between thresholds.

If orientation is purely a lenticular one, could these variations have been produced by effects of body-position on the lens? Since the targets were 10 ft. from *O*, the lens was not relaxed, and it has long been known that the accommodated (relaxed) lens is affected by gravity, while the unaccommodated lens is not.[25] In any event, this fact would not explain the uniform increase in threshold-difference in the body-tilt condition compared to the head-tilt condition when the position of the head was the same.

---

[24] Shortly after the start of this experiment, MS fell ill. When it was learned that he would not be available for a long time, the writer served as *O* in his place.

The question arises as to whether these results are due to counter-rotation of the eyes when the head is tilted. There has been considerable disagreement about the extent of such counter-rolling. For example, Adler states that "the vertical meridians of the cornea stay erect and do not incline with the head."[26] Scobee says, however, that

> The eyes were not capable of rotation more than 45° to 55° from the primary position. . . . Moreover, 99% of our ocular excursions are less than 15° . . . , from the primary position—a curiously overlooked but enormously important fact. To protect the optic and ciliary nerves and blood vessels from undue stretching or jarring when the eyes are rotated, there are the check ligaments.[27]

Indeed, the check ligaments apparently function very effectively. Woellner and Graybiel, in a study which has provided us with quantitative rather than descriptive information, have found that tilting the head 45° results in only a 3.5° counter-rolling of the eyes.[28] This would not be enough to account for the marked changes in acuity, since target-rotations of this magnitude result in only very small changes in acuity.[29] Nor would it, again, account for changes with the head in a constant position.

We conclude that the orientation is affected by extra-ocular factors, and that the appearance of the full effect probably depends on the absence of conflicting cues and an unequivocal conception of verticality (whether accurate or not). Interestingly enough, the phenomenon has also been found and studied in the octopus by Sutherland, who has attempted to explain it on a retinal basis.[30] Reid has found a movement-illusion analogous to the horizontal-vertical visual illusion.[31] If a non-visual illusion analogous to the orientation were found, it would be an interesting argument in favor of the psychological explanation of the effect. Finally, it should be noted that Hubel and Wiesel have found that the receptive fields in the cat's visual cortex respond to various types of stimuli only when the stimuli are presented with rather restricted orientations.[32] As yet, they have not found that fields sensitive to one orientation are more prevalent than those sensitive to another. If fields oriented to the vertical and horizontal were found to outnumber those oriented toward the diagonals, it again would be difficult to see why such a fundamental physiological phenomenon should be disturbed by the position of the body.

---

[25] C. Hess, Beobachtungen über den Akkommodationsvorgang, *Klin. Monatsbl. Augenh.*, 42, 1904, 309-315.

[26] F. H. Adler, *Physiology of the Eye*, 3rd. ed., 1959, 409.

[27] R. G. Scobee, *The Oculorotary Muscles*, 1947, 69.

[28] R. C. Woellner and A. Graybiel, Counterrolling of the eyes and its dependence on the magnitude of gravitational or inertial force acting laterally on the body, *J. appl. Physiol.*, 14, 1959, 632-634.

[29] Ogilvie and Taylor, *op. cit.*, *J. opt. Soc. Amer.*, 48, 1958, 628-629.

[30] N. S. Sutherland, Visual discrimination of orientation and shape by the octopus, *Nature*, 179, 1957, 11-13.

[31] R. L. Reid, An illusion of movement complementary to the horizontal-vertical illusion, *Quart. J. exp. Psychol.*, 6, 1954, 107-111.

Summary. Acuity-thresholds were determined for checkerboard targets in horizontal-vertical and diagonal orientations for four *O*s seated in three positions: (1) a normal upright position; (2) in an upright chair but with the head tilted; and (3) in a tilted chair. Cues for the true vertical were provided in the first two conditions, but false cues corresponding to the tilt of the head were given in the last condition. Thresholds were better for targets parallel to the median plane of *O*'s head, but the difference between the thresholds for the two orientations varied with *O*'s posture: the greatest difference occurred in the upright position, the smallest difference with only the head tilted 45°, and an intermediate difference with both body and head tilted. Tilting only 22.5° from the vertical produced no such uniform changes with the targets oriented at 0° and 45°. These changes were not the result of such factors as discomfort from the tilted chair, for no such changes occurred when the head was kept in an upright position while the chair was tilted. Support for the sensory-tonic theory was found in that acuity suffered under the unusual postures. It was concluded that the appearance of the full orientation-phenomenon depends on an unequivocal conception of verticality by *O*, whether accurate or not.

---

[32] D. H. Hubel and T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.*, 160, 1962, 106-154.

FORMATION AND GENERALIZATION OF CONCEPTS AS A FUNCTION OF SIMILARITY AND AMOUNT OF TRAINING

By R. DALE DICK, Fort Hays Kansas State College

While it has received considerable attention as a variable in other forms of learning, similarity has not been emphasized in the experimental study of conceptual behavior. The present study was concerned with the relation of similarity of forms to the rate of formation of concepts and amount of training to generalization of concepts.

From her verbal learning postulates, Gibson deduces: "If multiple generalization occurs during the learning of a list, and if the list is constituted so that the generalizing stimulus items are paired with the same responses, learning should be easier in proportion to the degree of generalization."[1] She defines generalization as contingent on similarity among stimuli, and points out that this situation is commonly employed in studying formation of concepts. Underwood says that generalization is an explanatory concept which provides the possibility of a rapprochement of conditioning, rote-learning, and tasks which can be catagorized as thinking.[2]

The first hypothesis was that the formation of concepts is faster for concepts of high similarity than for concepts of low similarity, 'high' and 'low' referring to the degree of similarity among instances or examples of a single concept—intra-concept similarity. Hull, Reed, and Welch have identified characteristics of concepts which facilitate concept-formation.[3] Their findings can be interpreted as due to similarity among the instances of concepts, but this interpretation is questionable. Hull may have been manipulating a perceptual rather than a learning variable; Reed and Welch may have been requiring their *Ss* to use conceptual responses they had acquired in every-day experience.

The second hypothesis is derived from Gibson's fourth postulate on intralist generalization, which she has confirmed:[4] "Generalization will increase to a maximum or peak during the early stages of practice with a list, after which it will decrease as practice is continued."[5] The conditions for generalization of paired associates and concepts are similar, the only difference being that a single stimulus is used in paired associates and a group of stimuli is used in concepts. If Gibson's postulate applies to the former, it may apply to the latter. Accordingly, the second hypothesis was that generalization of concepts (that is, the elicitation of a conceptual response by a stimulus which is not identical with any stimulus involved in formation of the concept) first increases as amount of training increases and then decreases as training continues.

---

* Received for publication January 3, 1962. This report, based on work submitted in partial fulfillment of requirements for the Ph.D. degree at the University of Missouri, 1958, was directed by Professor Robert S. Daniel. The work was supported in part by the United States Air Force under Contract AF 18(600)-1052.

[1] E. J. Gibson, A systematic application of the concepts of generalization and differentiation to verbal learning, *Psychol. Rev.*, 47, 1940, 208.

[2] B. J. Underwood, An orientation for research on thinking, *Psychol. Rev.*, 59, 1952, 216.

[3] C. L. Hull, Quantitative aspects of the evolution of concepts. An experimental study, *Psychol. Monogr.*, 28, 1920 (No. 123), 1-86; H. B. Reed, Factors influencing the learning and retention of concepts: I. The influence of set, *J. exp. Psychol.*, 36, 1946, 71-87; Livingston Welch, A preliminary investigation of some aspects of the hierarchical development of concepts, *J. gen. Psychol.*, 22, 1940, 359-378.

[4] Gibson, Intra-list generalization as a factor in verbal learning, *J. exp. Psychol.*, 30, 1942, 185-200.

[5] Gibson, *op. cit.*, 1940, 206.

It should be noted, however, that the decrease in generalization in the above postulate is derived from the assumption that differentiation among the generalized stimuli occurs in the later stages of training. This leaves open the question of generalization with stimuli which are not differentiated during training. Gibson's tenth postulate refers to such a situation: "If differentiation has been set up among a number of stimulus items, there will be less tendency for them to generalize with *new* stimulus items or for *new* stimulus items to generalize with them, the decrease in generalization being proportional to the amount of differential reinforcement given."[6]

METHOD

Semi-randomly generated visual forms (eight-sided, solid, flat figures) were used. Concepts (sets of visual forms) were constructed to show high similarity; others, to show low similarity. The first hypothesis was tested by comparing the rates of acquisition of conceptual responses for these two types of concept. The second hypothesis was tested by comparing the generalization of conceptual responses for four amounts of training.

Stimuli. The stimuli were organized as two related groups of forms, as shown in Fig. 1. The two groups were related by a series of successive deviations (the horizontal array in Fig. 1) through which the prototype (P_A in Fig. 1) of one group changed into the prototype (P_B in Fig. 1) of the other group. Each group consisted of concepts of high and low similarity (the upper and lower vertical arrays, respectively, in Fig. 1). The instances of both concepts were distortions along the same dimension of a single prototypical form. This made the schema of the two concepts the same. Thus, there were four concepts in the basic organization: a concept of low similarity and a concept of high similarity from each of two related prototypes.

This basic organization was developed five times, making a total of 10 pairs of concepts of high and low similarity. Five series of deviations related the prototype of one pair of concepts to the prototype of another pair.

Concepts. Each prototype was distorted to obtain the instances of the concepts of high and low similarity (the vertical arrays of Fig. 1) by tracing projections of the prototype on a screen set at various positions of rotation about its vertical axis. The units separating the five levels of distortion (left of Fig. 1) for the concepts of low similarity were twice as big as those for the concepts of high similarity. These units were specified in terms of the secant function for the angle of displacement of the screen from perpendicular to the line of projection. The tracings were made with the image at systematically varied positions of rotation about its center to get systematic variation of the axis of distortion or 'stretching' of the prototype. There were two such settings differing by 90° (the double columns in

---

[6] Gibson, *op. cit.*, 1940, 207 (italics mine).

Fig. 1) for each position of the screen. The maximal dimensions of the traced images were made approximately the same by adjusting the throw to compensate for changes in over-all size produced by the distortion. The names of the concepts were nonsense-syllables taken from Glaze's list of syllables with zero association value.[7]

The instances of each concept differed on a specified physical dimension (amount of distortion) and yet retained the same schema, in that they were obtained from a single prototype. Stevanović used a simple form of the method used here to generate such concepts.[8] There were also analogous concepts, having the same schema, at two levels of similarity which differed on a specified physical dimension (range of distortion).

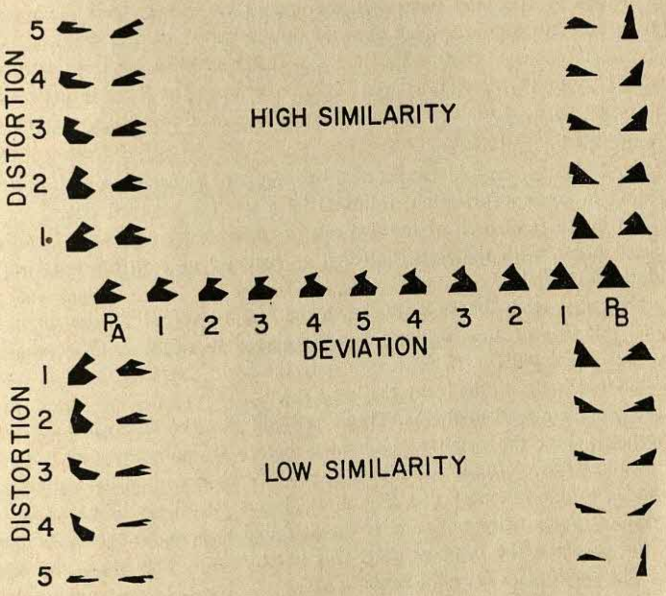


FIG. 1. BASIC ORGANIZATION OF THE STIMULI OBTAINED FROM A PAIR OF RANDOMLY GENERATED FORMS

Test-stimuli. The prototypes and the deviations were the test-stimuli. A modification of Method I of Attneave and Arnoult was used to generate these stimuli.[9] The series of deviations was obtained by systematically moving the points on which the figures were drawn. A series of forms similar to these was used by Mitsos.[10]

---

[7] J. A. Glaze, The association value of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 255-269.

[8] B. P. Stevanović, An experimental study of the mental processes involved in judgement, *Brit. J. Psychol. (Monogr. Suppl.)*, 12, 1927, 1-138.

[9] Fred Attneave and M. D. Arnoult, The quantitative study of shape and pattern perception, *Psychol. Bull.*, 53, 1956, 452-471.

[10] S. B. Mitsos, Permeability: A dimension of conceptual behavior, Unpublished Doctoral dissertation, Purdue University, 1955, 1-70.

The test-stimuli were deviations from the concepts on a specified physical dimension which differed from the dimension of similarity used in generating the concepts. They were progressive changes in schema, whereas the instances of each concept were distortions in which the schema was retained. They represented six levels of deviation from the concepts, with the prototypes at zero deviation and the forms in the middle of the series of deviations (Fig. 1) at Level 5.

Subjects. The *S*s were 160 volunteers (125 men and 35 women) from 10 beginning classes in psychology at the University of Missouri, run in groups of 2 to 17 with the exception of 5 who were run singly.

Apparatus. A Bell and Howell Robomatic projector equipped with a timer, an external shutter driven by a stallable motor, and an Industrial timer were used. These were arranged to provide for a continuously repetitive cycle of projection of only the upper part of a slide, containing an instance, for a controlled interval of time, followed by projection of the whole slide, showing the instance and its name, for a controlled interval of time.

Procedure. The *S*s received the set of 10 concepts of either high or low similarity for training in concept-formation, followed by a test for generalization.

The 100 instances of each of the two sets of concepts were presented in the same randomized order, with the restriction that an instance of each concept occurred in every block of 10 trials.

Before the first trial, the *S*s were shown the first 10 stimuli in the series for the controlled intervals of time without being required to respond. On each trial, an instance was presented for 10 sec., with only the form showing the first 7 sec. and the syllable together with the form, the last 3 sec.

Instructions were read to the *S*s. They were told that the task was to write down on a response-sheet the name of each form before its name appeared; that there would be 10 classes of forms and all the forms in each class had the same name and were similar to one another; that they must always attend to the entire form; and that they must guess when not sure. If the *S*s failed to respond before the name appeared, the space on the response-sheet was left blank. If they responded correctly, they placed a check mark after the response.

Upon completion of training, the *S*s were given a test consisting of presentation of 55 stimuli (the prototypes and deviations) without the syllables in random order for 10 sec. each. The task was to give the correct name for each form, *i.e.* the name of the concept from which the test-stimulus was a deviation. In the instructions, which were read, the *S*s were told that each form would be similar to one of the classes they had learned, and that they had to make a response to each form.

Throughout the procedure the names of the concepts were on a large poster beside the screen.

Design. A Type III design was used.[11] There were 8 independent groups of 20 *S*s each, 1 receiving concepts of low similarity and 1 receiving concepts of high similarity, at each of 4 amounts of training—presentation of the 100 instances 1, 2, 3, and 4 times, respectively. The groups will be referred to as HS and LS for high and low similarity, respectively; and as I, II, III, and IV, for the four amounts of training. The dependent variables in the design were levels of distortion in the data in training and the levels of deviation in data on the test for generalization.

---

[11] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education,* 1953, 281-284.

Each *S* was allowed to choose one out of 8 or 10 of the total of 23 experimental sessions. The groups of *S*s who chose the first 8 sessions were assigned at random to the 8 experimental groups. The *S*s who chose the remaining sessions were assigned to the experimental groups as necessary to equalize the *N*s for the groups as the sessions progressed.

Data from 10 *S*s were discarded. Four failed to follow instructions to guess and did not respond on more than 50 of the first 100 trials; one failed to understand instructions for the test; two did not complete the experiment; and three were considerably older than the other *S*s.

RESULTS

Training. The mean correct responses on the last 100 trials of training at each level of distortion are shown in Fig. 2. Since there were two instances for each of 10 concepts at each level of distortion, a maximum of 20 correct responses was possible at each level.

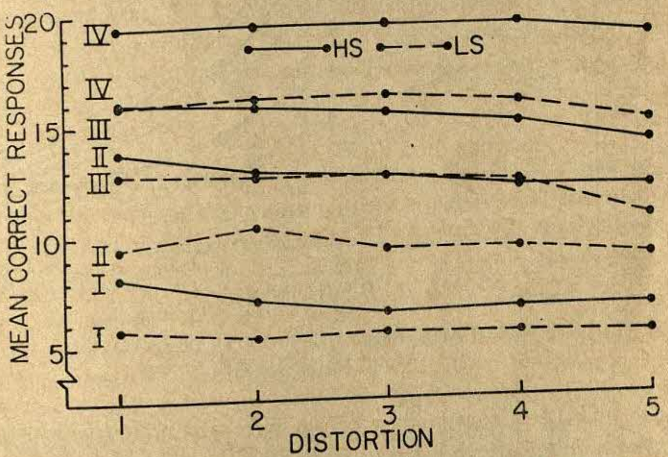


FIG. 2. MEAN CORRECT RESPONSES AT EACH LEVEL OF DISTORTION ON LAST 100 TRIALS OF CONCEPT-FORMATION

Analysis of variance of these data yielded the following significant effects: Training ($F = 86.29$, $df = 3/152$, $p < 0.01$), Similarity ($F = 29.47$, $df = 1/152$, $p < 0.01$), Distortion ($F = 14.67$, $df = 4/608$, $p < 0.01$), and Distortion × Training ($F = 2.77$, $df = 12/608$, $p < 0.01$). Both error-variances were heterogeneous, but the 1% level of significance can be safely interpreted as indicating significance at least at the 5% level.[12]

The Tukey Gap Test was applied to the mean differences pertinent to the hypothesis.[13] The means for amounts of training for HS were 6.93, 12.75, 15.37, and 19.49, for I through IV; for HS they were 5.47, 9.55, 12.25, and 15.96. All the differences between the means for successive amounts of training, separately for HS and LS, were significant (smallest significant difference = 2.72, $df = 152$, $p < 0.01$), except II vs. III for both HS and LS. Conceptual responses were an increasing function of amount of training at both levels of similarity.

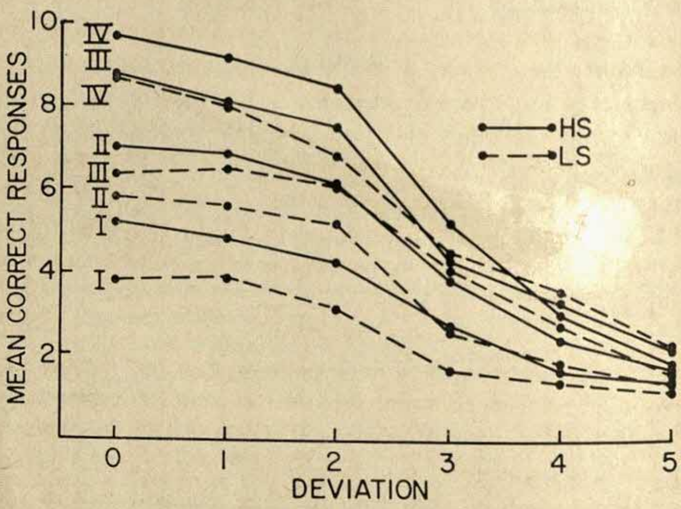


FIG. 3. MEAN RESPONSES AT EACH LEVEL OF DEVIATION ON TEST OF GENERALIZATION

The differences between the HS and LS means were significant for II, III, and IV, but not for I. Acquisition of conceptual responses was faster for HS than for LS.

Test. The mean correct responses at each level of deviation are presented in Fig. 3. Since there was one deviation for each of 10 concepts at each of the first five levels of deviation, a maximum of 10 correct responses was possible at each of these levels. Responses at the sixth (middle) level of deviation were counted as correct if they were the name of either of the concepts of which the form was a deviation. Since there were five stimuli at this level, a maximum of five correct responses was possible and one correct response was possible by chance, as was the case at the other levels of deviation.

---

[12] Lindquist, *op. cit.*, 86.

[13] J. W. Tukey, Comparing individual means in the analysis of variance. *Biometrics*, 5, 1949, 99-114.

In the analysis of variance, the following sources were significant: Training ($F = 30.69$, $df = 3/152$, $p < 0.01$), Similarity ($F = 13.28$, $df = 1$ and 152, $p < 0.01$), Distortion ($F = 600.19$, $df = 5/760$, $p < 0.01$), Distortion × Training ($F = 11.74$, $df = 15/760$, $p < 0.01$), and Distortion × Similarity ($F = 8.60$, $df = 5/760$, $p < 0.01$). The between variances were not found to be heterogeneous, while the within variances were heterogeneous. Tests of significance based on the latter error-variances were made at the 1% level and can be interpreted as significant at at least the 5% level.[14]

The means for amounts of training for HS were 3.31, 6.63, 5.63, and 6.19 for I through IV; for LS they were 2.51, 3.72, 4.52, and 5.63. Mean differences between the successive amounts of training were all significant ($p < 0.01$) except those for HS-III and IV and LS-II and III.[15] Conceptual generalization thus tended to be an increasing function of amount of training.

DISCUSSION

The prediction from the first hypothesis that the number of correct responses in the last 100 trials would be greater for HS than for LS was verified for amounts of training II, III, and IV. This effect of similarity has been accounted for by Gibson as due to stimulus-generalization among the instances of each concept.[16] The generalization of conceptual responses within each concept was greater for concepts of high similarity than for concepts of low similarity.

The prediction from the second hypothesis was not supported. In her description of verbal learning, Gibson assumed that both generalization and differentiation occur before the learning process is complete.[17] The various groups in the present data range from the early stages (LS-I) to nearly complete concept-formation (HS-IV) as shown in Fig. 2 and by the significant differences among the means for successive amounts of training. If Gibson's fourth postulate applies to generalization of concepts as well as to generalization of paired associates, the test-scores should have decreased at the greater amounts of training; but generalization of concepts appears as an increasing function of amount of training.

---

[14] Lindquist, *op. cit.*, 86.

[15] Tukey, *op. cit.*, 99-114.

[16] Gibson, *op. cit.*, 1940, 196-229.

[17] Gibson, *op. cit.*, 1940, 196-229.

It will be remembered that differentiation of the stimuli, on which generalization was tested, was not required during training in the present experiment, whereas such differentiation is assumed in Gibson's fourth postulate. In any event the present finding is contrary to Gibson's tenth postulate, at least as it applies to generalization to new stimuli which are not differentiated during training. Differentiation among concepts does not appear to affect a decrease in the generalization of conceptual responses to new stimuli that are similar to the concepts.

The use of generalization as an explanatory concept for a rapprochement of conditioning, rote-learning, and tasks catagorized as thinking, suggested by Underwood,[18] may be questioned. Hovland found that generalization of conditioned galvanic skin responses at first increased and then decreased as training continued from 8 to 12 to 24 reinforcements, without differentiation of the generalized stimuli.[19] Gibson reported a similar finding for generalization within lists of paired associates, but the generalized stimuli were differentiated.[20] Generalization of concepts may be a different function of training than generalization of conditioned responses or paired associates.

The effects of generalization obtained here are limited to responses acquired to points along one dimension of similarity generalizing along a *different* dimension of similarity. Conceptual responses acquired on one segment of a dimension might generalize to another segment of the *same* dimension in a manner quite different from the effects of generalization obtained here.

If further research using different conditions for generalization is expected to yield results which are similar to those of the present study, the following conclusions would be predicted: The rate of formation of concepts is faster with concepts of high similarity than with concepts of low similarity. Generalization of concepts is an increasing function of amount of training when differentiation among the generalized stimuli does not take place during training.

SUMMARY

Eight groups of *S*s were each given one of four amounts of training in concept-formation with concepts of either high or low similarity. Each of 10 random, flat, solid, visual forms was the prototype for 10 instances

---

[18] Underwood, *op. cit.*, 216.

[19] C. I. Hovland, Generalization of conditioned responses, *J. exp. Psychol.*, 21, 1937, 261-276.

[20] Gibson, *op. cit.*, 1942, 185-200.

of each of 10 concepts of high, and 10 concepts of low, similarity. Instances of concepts of low similarity represented a greater range of distortion, achieved by projection of the prototype on a rotated screen, than instances of concepts of high similarity. After training, the *S*s were given a test for generalization, consisting of stimuli obtained by systematically moving the points on which the prototypes were drawn.

It was found that the rate of concept-formation was faster with concepts of high than with concepts of low similarity and that generalization of concepts was an increasing function of amount of training. These findings were interpreted in relation to Gibson's theory of verbal learning.

COGNITIVE AND S–R INTERPRETATIONS OF INCENTIVE-MOTIVATIONAL PHENOMENA

By J. J. FUREDY and R. A. CHAMPION,
University of Sydney, Australia

Following Bitterman's delineation of opposing predictions derivable from the theories of Spence and Tolman,[1] Gonzalez and Diamond obtained experimental results which were interpreted as putting the differences between the two theories in clearer perspective.[2] Their most striking finding concerned two groups of rats given a series of unrewarded trials in a runway, *Rx,* which was dissimilar to the empty end-box, *By.* In a second stage, the members of one of these groups (Group B) were directly placed and fed in *Bx,* a box similar to *Rx* but unlike *By,* while the other rats (Group C) were fed in *By.* It was deduced that, in a third stage of further unrewarded trials in *Rx By,* Spence should predict faster running in Group B than in Group C, whereas Tolman should predict the opposite result; the data were in accordance with the prediction attributed to Tolman. While, however, Bitterman's deductions follow directly from Spence's theory as explicitly stated,[3] there remains the possibility that Spence did not refer to some factors which might also be taken into account in any further test of S–R theory, and the following study was designed to allow for the possible role of two such factors.

In outline, Spence has supposed that when a goal-object such as food leads to a goal response (R_g) in the end-box of a runway, cues (S_2) temporally contiguous with R_g come to elicit some component (r_g) of this response through classical conditioning; if stimuli in the starting box and runway (S_1) bear any similarity to S_2, then they may elicit r_g anticipatorily through stimulus-generalization, so that the stimulus-properties of r_g (s_g) become part of the complex of stimuli eliciting the instrumental response of running. By assuming that the intensity of s_g varies with the frequency and vigor of R_g, Spence has envisaged a mechanism mediating the known activating effects of food-reward. It is not a necessary feature of this model, however, that s_g should come to elicit the particular instrumental response under scrutiny by an experimenter, for if some response other than running should be dominant in the runway-situation, then s_g could well come to form part of the complex eliciting this competing response and so interfere with the response being

---

* Received for publication December 7, 1961.

[1] M. E. Bitterman, Review of Spence's *Behavior Theory and Conditioning,* this JOURNAL, 70, 1957, 141-145, esp. 144.

[2] R. C. Gonzalez and Leonard Diamond, A test of Spence's theory of incentive-motivation, this JOURNAL, 73, 1960, 396-403.

[3] K. W. Spence, *Behavior Theory and Conditioning,* 1956, 134-136.

measured. It is emphasized that, according to this elaboration of Spence's hypothesis, it is not the hypothetical r_g but some overt response which is thought to compete with the instrumental response under study. The experiment of Gonzalez and Diamond began with a series of unrewarded trials in a 48-in. runway, the running times being of the order of 40 sec., and these times increased slightly over the first set of trials, suggesting that competing responses were strong. This, therefore, was the first factor which the present study sought to control, by the use of rewarded trials in the first stage so as to ensure the clear dominance of the running response.

To make a complete application of Tolman's type of theory to the data of their study, Gonzalez and Diamond extended the theory by supposing that direct-placement feedings in *Bx*, for example, acted not only to increase the valence of the stimuli in *Bx* as an end-box (S_2) but also the valence of the cues in *Rx* (S_1), because of their similarity to S_2. Thus the performance of Group B should have been interfered with by direct feedings in *Bx* because of the similarity of *Bx* to *Rx* and the dissimilarity of *Bx* to *By*, making the valence of the stimuli in *By* (S_2) relatively less than that of the stimuli in *Rx* (S_1). It can be shown, on the other hand, that S–R theory leads to such an expectation without any extension. According to Spence's hypothesis, the direct-placement feedings in *Bx* or *By* cause r_g to be elicited by the stimuli there (S'_2) so that, other things being equal, r_g should occur more strongly in *Rx* with feedings in *Bx* because of the greater similarity between *Bx* and *Rx* than between *By* and *Rx*. But again the model does not require that only r_g is conditionable to S'_2, and it must be recognized that under the somewhat unusual direct-placement conditions the cues in *Bx* and *By* may come to elicit a variety of other overt responses which might be loosely characterized as 'looking for food,' that these will transfer to *Rx* more from *Bx* than from *By*, and that performance could well be inferior in Group B for this reason. This transfer could also occur under normal runway-training conditions, as Hull envisaged,[4] but the competing responses are assumed to become subordinate to the reinforced instrumental response of running. The r_g, on the other hand, is also transferred, but is thought of as incapable of competing with any overt response. To make some allowance for this second factor which emerged from a more detailed consideration of S–R theory, an attempt was made to reduce competing, overt ('looking-for-food') responses to a minimum by sounding a tone whenever *S* entered a goal-box containing food (*Bx* or *By*), this stimulus not being presented while *S* was in *Rx*. In Skinner's terms, it was intended that the tone should become a discriminative stimulus for food. The use of this stimulus should not have destroyed the rationale of the experiment, the similarity between *Bx* and *Rx* still being greater than that between *By* and *Rx* (Spence), with the relative valence of *By* still greater than that of *Bx* (Tolman).

Method: (1) *Subjects.* The *Ss* were 16 experimentally naïve male albino rats, ranging in age from 83 to 107 days at the beginning of preliminary training.

(2) *Apparatus.* The apparatus was similar in function to that used by Gonzalez and Diamond, but it differed in structural detail. There were two runways, 5 in. wide, 24 in. long, and 5½ in. high. One runway had white walls and a wire-

---

[4] C. L. Hull, *A Behavior System*, 1952, 124-125.

covered aluminum floor, while the second runway had black walls and an uncovered floor. There were also four boxes, 5 × 6½ × 5½ in., which could be used interchangeably as starting boxes or end-boxes, two similar to the white runway and two similar to the black runway. Another set of four boxes, identical with the first set except for their greater length of 9½ in., was used in the second replication because the *S*s were older and larger. The runways and boxes were covered with clear perspex. The boxes were separated from the runways by means of transparent guillotine-doors, while the interior of each goal-box was hidden from the runway by an opaque curtain on the runway-side of each door. When *E* lifted the door between the starting box and the runway a micro-switch was automatically closed so as to start an electronic chronoscope, and the passage of the rat's body across a gap in the metal floor between starting box and runway activated a relay which stopped the chronoscope, thus giving a measure of starting time. Running times were measured less precisely, with a stopwatch. Another microswitch was linked with the door between the runway and the goal-box, so that when *E* closed this door as *S* entered the goal-box, a 500-cpsec. 85-db (spl) tone was sounded for 5 sec.

(3) *Procedure.* The experiment was conducted in two replications, with 6 *S*s in the first and 10 *S*s in the second.

(4) *Preliminary training.* For the first 10 days the *S*s were put on a 24-hr. feeding schedule, handled for 5 min. daily, and allowed to eat for 1 hr. On the eighth day the food was changed from dry pellets to wet mash. An equal number of *S*s was assigned to the white and black runways, and they were allowed to explore the apparatus with the doors open, no food-reward, and with no tone, for 5 min. daily until all *S*s were readily crossing into the goal-box through the curtain. This pretraining was conducted immediately prior to the daily eating session; it took three days for the first replication and six days for the second replication.

Stage 1. After at least 2 min. of adaptation in the experimental room, each *S* was given training trials in the appropriate runway, as follows. When *S* oriented towards the door of the starting-box, the door was opened by *E* and then lowered after *S* had passed into the runway, so as to prevent retracing. When *S* passed through the curtain into the goal-box, the second door was lowered, to prevent retracing, and the tone sounded for 5 sec. The reward for each trial was approximately 2 gm. of wet mash, half food and half water by weight. As soon as *S* finished eating (after approximately 1 min.), it was replaced in the starting box for the next trial. The *S*s were given 4 trials on each of two days, being allowed to eat wet mash ad lib for 30 min., starting at least 30 min. after the 4th trial each day.

Stage 2. On the basis of mean starting times over the last 4 trials of Stage 1, the two main Groups B and C were established by matching the *S*s in pairs, with the restriction that half the *S*s in each group be trained in the white and half in the black runway. Then the *S*s were given direct-placement feedings, either in a box similar to the starting box and runway of Stage 1 (*Bx,* Group B) or in the goal-box of Stage 1 (*By,* Group C). After 3 min. of adaptation in the experimental room each *S* was placed directly into the box containing approximately 4 gm. of wet mash (50% dry food by weight) and the tone was sounded for 5 sec. As soon as *S* had finished eating, it was returned to the holding cage for 1 min. and then replaced in the box for a further period of eating. The *Ss* were fed four times on each of four days in this fashion, and once on the fifth day. On the first four days the *Ss* were allowed to eat mash ad lib for 30 min., beginning at least 30 min. after the last direct-placement feeding, but on the fifth day the single feeding in the box was followed by the first four trials of Stage 3. By making small adjustments in the amount of mash presented in the box over the 17 trials of Stage 2, it was ensured that the average time spent in eating in the goal box per trial by each *S* was approximately 120 sec. (ranging from 117 to 126 sec.) and that the average times for Groups B and C were equal at 120.7 sec. As already noted, care was also taken to see that the entire time in the goal-box was spent in eating, in both Stage 1 and Stage 2.

Stage 3. After the direct-placement feeding on the fifth day of Stage 2, all *Ss* were removed from the experimental room. At least 15 min. later they were returned to the room, allowed 2 min. of adaptation, and then given 4 trials in the manner of Stage 1. A further 4 trials were given on the next day.

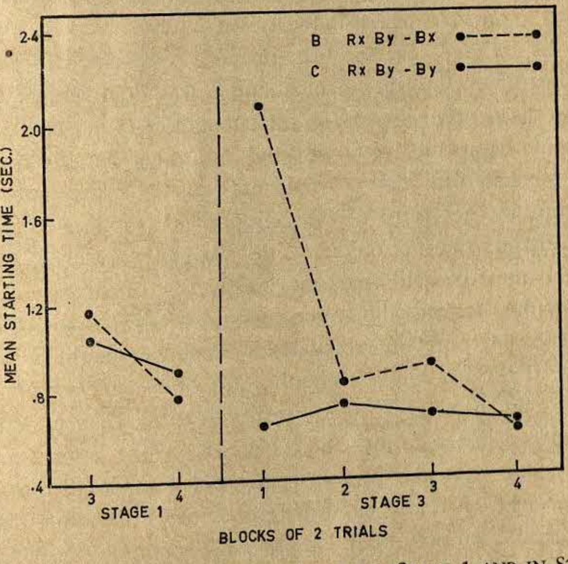


FIG. 1. PERFORMANCE IN THE SECOND HALF OF STAGE 1 AND IN STAGE 3

Results. The performance of the two Groups B and C on the last 4 trials of Stage 1 and the 8 trials of Stage 3 is plotted in Fig. 1, in terms of mean starting-time. The running times showed the same trends as a measure of performance, but were more variable than the starting-times. Taking the first block of two trials in Stage 3 as a measure of the effects of Stage 2, the individual mean starting-times (sec.) were as follows. Group B: 4.00, 2.50, 1.95, 1.80, 1.65, 1.60, 1.60, and 1.45; Group C:

0.85, 0.75, 0.70, 0.65, 0.60, 0.58, 0.50, and 0.50. Since there was no overlap in these two samples, no statistical test was deemed necessary, the performance of Group C being clearly superior. It may be observed in Fig. 1, however, that the difference between the two groups had all but disappeared by the second block of two trials in Stage 3.

Discussion. The chief impression provided by the data of this experiment is that of further support for Tolman's position, in view of the clear superiority of Group C in the early trials of Stage 3. It is generally accepted, however, that a theory is rarely demolished by the results of a single experiment, as a crucial instance, but rather tends to give way gradually to a superior theory. Since S–R theory appears to offer the most potent alternative to a cognitive interpretation of the data, the remainder of this discussion is devoted to the chief comments still open to the S–R theorist and to possible defects in the present study.

It will be recalled that the experiment was designed to ensure: first, that the instrumental response of running was dominant at the end of Stage 1, so that further increase in the strength of r_g in Stage 2 would not activate some competing response; and secondly, that 'looking-for-food' responses generated in Stage 2 would not transfer to the starting box and so constitute another form of competition with running as measured in Stage 3. Whereas the first of these aims was easily achieved, it is felt that some doubts must be held about the effectiveness of the tone as a discriminative stimulus. Indeed, the search for an effective means of preventing the generation of competing responses in Stage 2 leads the S–R theorist to a dilemma. If the feeding-box stimuli (S'_2) of Stage 2 and the stimuli of starting box and runway (S_1) are made identical, then although r_g may be strengthened in the former and transferred to the latter situation, competing responses may also be transferred, being made all the stronger by the dynamogenic effects of r_g. If S'_2 and S_1 are made quite dissimilar on the other hand, then not only the undesirable competing responses but also the activating r_g may fail to transfer. In the present experiment a compromise was sought by presenting the tone in S'_2 (feeding box) and S_2 (goal-box) only for the first 5 sec., while the animal was moving towards food, and by not sounding the tone for any length of time during actual eating, when r_g was presumably occurring. If this technique failed in its purpose, however, then the performance of Group B would have been depressed in comparison with that of Group C, because of the greater similarity of *Rx* to *Bx* than to *By*, and this was the obtained result.

In examining the results of the present study, attention was concentrated upon the comparison of Groups B and C within Stage 3. This was done in the belief that comparisons between levels of performance at different stages have no relevance to the comparison of the theories of Tolman and Spence. It might be suggested that some meaning could be given to between-stage comparisons by the use of a control group of *S*s given no direct-placement feedings in Stage 2 but simply held in their home cages and given normal daily feedings. The irrelevance of between-stage comparisons and the ineffectiveness of this control group may be demonstrated by consideration of the following possibilities in terms of Fig. 1. First, taking the data at face-value, the performance of Group C was relatively unaffected by the events of Stage 2 whereas that of Group B was considerably impaired. The comparable result for Group C in the study of Gonzalez and Diamond was a marked increase in response strength, and the failure to obtain such an effect in the present case may have been due to the use of reinforced trials in Stage 1, which alone could have brought performance near a maximum. It is not usual for 8 trials to suffice for this purpose in runway-training, but further evidence for the possibility is to be found in the fact that Group C showed no further improvement in the course of Stage 3. As far as Group B is concerned, the modified cognitive theory would account for the decrease in response-strength by reference to the superior valence of the runway (*Rx*) relative to the goal-box (*By*) as a result of feedings in a box similar to the runway (*Bx*) in Stage 2. An S–R theorist, on the other hand, could suggest that the tone failed to act as a discriminative stimulus, so that the direct-placement feedings in *Bx* simply led to the occurrence of competing responses in *Rx* in Stage 3. It could even be suggested that the use of the tone was ill-advised in that, as a result of differential conditioning, it prevented transfer of r_g to the starting box and runway by its absence there. The answer to this claim, however, would seem to lie in the substantial improvement in performance which occurred in Stage 1, under conditions where the tone was present briefly in the goal-box but never in the runway.

Instead of taking the data at face-value, it might be allowed that some forgetting would occur in a group which did not have any direct feedings in Stage 2, enough to locate such a control group somewhere between Groups B and C at the beginning of Stage 3. If this were the case, the cognitive and S–R theories would seem to apply equally well to the data, for both could predict improvement in Group C and impairment of performance in Group B as a result of the direct-placement feedings. Assuming that the position of the cognitive theorist has already been made clear, the reason for such an expectation on the part of the S–R theorist may need elaboration. If the reinforced trials of Stage 1 are to have any effect then it must be through the r_g–s_g mechanism as outlined above; *i.e.* the r_g comes to occur anticipatorily in *Rx* through stimulus-generalization from *By*, for although *Rx* and *By* differ in important respects they still bear some degree of similarity to each other. This being so, direct feedings in *By* should lead to some further improvement in performance in Stage 3, provided that performance has not already reached a maximum by the end of Stage 1 or allowing that some forgetting would have occurred otherwise, after the end of Stage 1. It is to be presumed that the feedings in *By* would not cause competing responses in *Rx*, for otherwise the learning in Stage 1 would not occur so readily. In the case of Group B, on the other hand, the S–R theorist may argue as for the face-value case.

If this line of argument were followed to the limit it is conceivable that a control group denied feedings in Stage 2 could deteriorate in performance to the same extent as Group B did in the present study (Fig. 1). Under these conditions it would be concluded that the interpolated feedings in *Bx* had no effect, whereas feedings in *By* led to considerable improvement in Group C, as found by Gonzalez and Diamond. Even so, the S–R theorist could suggest that feedings in *Bx* produced opposed factors which counter balanced each other in Group B, while the activating effects of r_g were free to show in the performance of Group C. It should be emphasized at this point that Spence's statement of his version of the S–R position does not explicitly refer to these theory-saving possibilities, and that the critics of this position have proceeded along perfectly proper lines in their interpretations of the theory as stated. Nevertheless, it is felt that an objective assessment of the position should exhaust every possibility.

In the light of these comments arising from a consideration of the present data, some general findings and interpretations of Gonzalez and Diamond may be reëxamined briefly. These *Es* have suggested that, according to S–R theory, the performance of groups fed in a box (*Bx*) similar to the starting box and runway (*Rx*) (their Groups A and B) should be superior to that of groups fed in a box (*By*) dissimilar to the rest of the alley (their Groups C and D). Now whereas this type of interpretation may be applicable to the usual runway-training situation, in a comparison of groups trained with *Rx Bx* and *Rx By*, it may be rendered invalid in the less usual, direct-placement situation by the generation of competing ('looking-for-food') responses in *Rx* with placements in *Bx*. Although this possibility was not allowed for by Spence in his brief outline of an experiment,[5] and not recognized at the outset by Stein when he conducted the experiment,[6] it need not be precluded from serious consideration. A second implication of S–R theory for Gonzalez and Diamond was its failure to account for effects obtained with variations in the similarity of feeding box (S'_2) to goal-box (S_2) with the relationship of S'_2 to starting box and runway (S_1) held constant. It is suggested, however, that their experiment has not produced clear evidence of the existence of such effects, in the form of statistically significant differences in Stage 3 between Groups A and B, and between Groups C and D. Finally, it bears repetition that the cognitive interpretation derived from Tolman completely encompasses all the findings.

SUMMARY

An experiment was conducted which sought to check the results of two groups used in an earlier study of Gonzalez and Diamond. These were

---

[5] Spence, *op. cit.*, 147-148.

[6] Larry Stein, The classical conditioning of the consummatory response as a determinant of instrumental performance, *J. comp. physiol. Psychol.*, 50, 1957, 269-278.

groups given a preliminary test without reinforcement in straight-alley apparatus with a goal box (*By*) dissimilar to the starting box and runway (*Rx*) and then given direct-placement feedings in a box (*Bx*) similar to *Rx* (Group B) or in *By* (Group C). The reason for this replication was the attempt to allow for the possible non-dominance of the instrumental, running response and for competing responses generated in the direct feedings in *Bx* and generalizing to *Rx* in a later test. Allowance was made for these factors by using reinforced trials before the direct feedings and by sounding a tone with all entries into *Bx* or *By*, as a discriminative stimulus for food. The result (Group C superior to Group B) was in keeping with a cognitive interpretation of the phenomena (Tolman), as proposed by Gonzalez and Diamond, but further features of an S–R interpretation (Spence) were considered and not excluded.

FURTHER STUDY OF THE RELATION BETWEEN THE SOUND OF A WORD AND ITS MEANING

By JONATHAN H. WEISS, University of Rochester

The phenomenon of phonetic symbolism is not a new or even a recently discovered one. Going back at least as far as Plato's *Kratylos,*[1] one finds it written that certain words seem to contain sounds which are in an intriguing manner appropriate or "fitting" to the meanings they embody. Thus, *scrape* or *slip* seem to display what may be termed high 'phonetic symbolism.' Just what role such a sound-sense relationship plays in language, however, has been problematical.

Several suggestions have been made, with one of which we shall be concerned here. Brown and Nuttall have offered the notion that languages evolve toward a state of richer phonetic symbolism, since words whose sounds are 'fitting' would be more easily retained and hence aided to survive.[2] While no direct evidence is available relating to this hypothesis, Wertheimer has reported a finding of some relevance.[3] In his study, it appeared that, in some instances, at least, English words which were high in 'fittingness' were also likely to be high in perceived intensity of meaning as measured by semantic differential ratings. Thus, words judged to have 'high fit' elicited somewhat more extreme ratings than 'low fit' words.

In attempting to understand this result, especially in relation to Brown's hypothesis, one is faced with an extremely interesting question: Is the demonstrated relationship due, at least in part, to some contribution of phonetic symbolism to the perceived meaning intensity, *e.g.*, a tendency to reinforce meaning characteristics, such that one could argue in support of Brown's notion that phonetic symbolism does have a discernible functional role in languages? Or may it be, on the other hand, as Jespersen[4] or Sapir[5] might have argued, that when words, which are essentially vocal gestures, are created, greater gestural participation is likely to accompany highly charged expressions, so that words which are highly intense would also be likely to be richer in phonetic symbolism; assuming, as Sapir does, that phonetic symbolism is the product of appropriate vocal gestures?

The present paper describes an attempt to shed some light on this question by assessing the effects of high vs. low 'fit' on the perceived in-

---

* Received for publication September 14, 1962.

[1] Benjamin Jowett (ed.), *The Dialogues of Plato*, 1892, 253-389.

[2] Roger Brown and Ronald Nuttall, Method in phonetic symbolism experiments, *J. abnorm. soc. Psychol.*, 3, 1959, 441-445.

[3] Michael Wertheimer, The relation between the sound of a word and its meaning, this JOURNAL, 71, 1958, 412-415.

[4] Otto Jespersen, *Language, Its Nature, Development, and Origin*, 1922, 421-442.

[5] Edward Sapir, A study in phonetic symbolism, *J. exp. Psychol.*, 12, 1929, 225-239, esp. 225.

tensity of meanings in words which are otherwise matched on an intensity dimension, viz. emotionality. The following hypothesis was tested: If phonetic symbolism reinforces the perceived meaning intensity of those words in which it occurs, then words which are judged to be low in 'fit' should elicit less extreme ratings on a measure of intensity, viz. the semantic differential, than words which are of equal emotionality but high in 'fit.'

Method and procedure: (1) Choice of words. From the high frequency words in the Thorndike-Lorge list, 127 monosyllabic words of no less than 3 nor more than 6 letters were chosen by *E*, randomized, and presented to two groups of judges.[6] The first group, 23 men and women undergraduate students in introductory psychology, was instructed to judge each word for sound-meaning 'fit' on a 3-point scale. The second group, comprised of 20 men and women from the same population, was instructed to judge the words for 'emotionality' defined as "the relative presence or absence of emotional connotations or overtones."

TABLE I

WORDS AND SEMANTIC DIFFERENTIAL SCALES

	Words	
	Hi-E	*Lo-E*
Hi-F	Howl	Tap
	Doom	Droop
	Whip	Glide
	Groan	Sweep
	Roar	Scratch
Lo-F	Bride	Root
	Wed	Rust
	Sin	Reach
	Vote	Route
	Steal	Barn

Factor	Scales	
Evaluation:	Good	—Bad
	Wise	—Foolish
	Kind	—Cruel
	Beautiful	—Ugly
Potency	Hard	—Soft
	Masculine	—Feminine
Activity	Active	—Passive
	Fast	—Slow
Novelty	New	—Old
Tautness	Angular	—Round

On the basis of these judgments 43 words constituting 4 preliminary groups, 2 of which were high in 'fit' and either high or low in emotionality (*Hi-F, Hi-E; Hi-F, Lo-E*) and the other 2 of which were low in 'fit' and either high or low in emotionality (*Lo-F, Hi-E; Lo-F, Lo-E*), were extracted, radomized, and submitted to 2 new groups of judges, numbering 8 and 9 respectively, for 'fit' and 'emotionality' ratings. Using these judgments, *E* extracted the 5 words best fitting the criteria for each of the 4 categories (see Table I).

(2) Choice of semantic differential scales. The scales used in this study were chosen from the table of scales and factor loadings given by Jenkins, Russell, and Suci.[7] The Evaluative factor was represented by 4 scales; Potency and Activity by

---

[6] E. L. Thorndike, and Irving Lorge, *The Teachers Word Book of 30,000 Words*, 1944.

[7] J. J. Jenkins, W. A. Russell, and G. J. Suci, An atlas of semantic profiles for 360 words, this JOURNAL, 71, 1958, 688-699.

2 each; and Novelty and Tautness by 1 each; giving a total of 10 (see Table I). The scales chosen were, in all cases, those most highly loaded on the factor represented.

(3) Administration. For each of the 20 words to be rated, a separate randomization of scales was used. In addition, the order of presentation of words was different for each *S*. The *Ss* were tested in a group. Each *S* was given a booklet made up of a cover sheet on which he filled out basic identifying information; an instructional sheet explaining and illustrating the use of semantic differential scales; and then 20 pages on each of which appeared 1 word and 10 scales.

(4) Subjects. Forty-five men and women, undergraduate students in introductory psychology who ranged in age from 18 through 23 yr., served in the present study. None of the *Ss* had previously served as a judge in the preliminary phases. All were native Americans who spoke English as their basic language.

Results. 'Extremeness of ratings' was defined in the present study basically as it had been by Wertheimer, viz. as the degree to which a concept was rated as deviating from the neutral point on a series of bipolar adjective scales. One important difference, however, exists between our measure of deviation and Wertheimer's, viz. where we dealt with discrete units on a 7-point scale with '4' representing neutrality, Wertheimer's technique made use of a continuous scale with *Ss*' being free to check any point on a straight line connecting the two adjectives. We thought that the more usual method employed in the present study is advantageous in that it increases the probability that scale units are uniform for all *Ss*, while Wertheimer's method allows for undetectable, idiosyncratic differences in the *Ss*' perception of scale units.

An 'extremeness' score for each of the four word-categories was derived by calculating *Ss*' mean deviations for all five words in each category. A separate score was calculated for each of the five factors represented. The scores thus derived were analyzed by means of Friedman's "Two Way Analysis of Variance by Ranks."[8] A summary of the analysis is presented in Table II.

An examination of Table II discloses two major findings. First, where the effects of *Hi-E* vs *Lo-E* are concerned, on three of the five factors (Evaluation, Potency, Novelty) *Hi-E* words, as expected, elicited more extreme ratings, although in only two of these instances (Evaluation and Potency), are the differences clearly significant. On the Tautness factor, the difference is reversed but non-significant, and on the Activity factor, mean scores are identical. Secondly, where the effects of *Hi-F* vs *Lo-F* were involved, on the Evaluation and Novelty factors, *Lo-F* words elicit significantly more extreme ratings, contrary to what one might be led to expect by Wertheimer's findings. On the Potency and Activity factors, however, *Hi-F* words do elicit more significantly extreme ratings in apparent accord with the hypothesis. These findings seem contradictory until one examines the likely contributing influences more closely. Thus, when one looks at the words in the *Hi-F* and *Lo-F* categories, it is immediately clear that five of the *Lo-F* words (bride, root, rust, barn, route) are nouns, while all *Hi-F* words are verbs. It is appropriate then that the Potency and Activity scales, which are sensitive to the action qualities of the concepts, should be checked at more extreme points when action words, *i.e.* verbs, predominate. The Evaluation and Novelty scales, on the other hand, are equally applicable to both nouns and verbs and thus show mainly the effect of emotionality and, if anything, the puzzling reversal, or "interference effect," due to the presence of fittingness which appears to depress Evaluation and Novelty ratings for *Hi-E* words. Evidence for this interpretation derives

---

[8] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences,* 1956, 166.

TABLE II

MEAN DEVIATION OF WORD GROUPS FROM NEUTRALITY ON SEMANTIC DIFFERENTIAL FACTORS

Evaluation

	Hi-E	*Lo-E*	Mean	
Hi-F	1.37	0.86	1.12	$P<0.05$
Lo-F	1.86	0.94	1.40	
Mean	1.61	0.90		
	$P<.05$			

Potency

	Hi-E	*Lo-E*	Mean	
Hi-F	1.44	1.12	1.28	$P<0.05$
Lo-F	1.31	0.96	1.13	
Mean	1.37	1.04		
	$P<.05$			

Activity

	Hi-E	*Lo-E*	Mean	
Hi-F	1.60	1.61	1.60	$P<0.05$
Lo-F	1.31	1.31	1.31	
Mean	1.45	1.45		
	N.S.			

Novelty

	Hi-E	*Lo-E*	Mean	
Hi-F	0.77	0.90	0.83	$P<0.05$
Lo-F	1.59	1.29	1.44	
Mean	1.18	1.09		
	N.S.			

Tautness

	Hi-E	*Lo-E*	Mean	
Hi-F	0.93	1.24	1.08	*NS*
Lo-F	0.95	1.00	0.97	
Mean	0.94	1.12		
	NS			

from an examination of mean deviations for verbs and nouns on the Potency-Activity and the Evaluation-Novelty scales (see Table III). It can be seen that on those factors (Potency and Activity) sensitive-to-action-qualities, verbs elicit more extreme ratings than do nouns, whereas on the Evaluation and Novelty scales, verbs and nouns elicit ratings which are not strikingly different.

On the whole, then, one must tentatively reject the hypothesis tested in favor of an alternative position which will be discussed below. The extremeness of ratings is thus apparently influenced for the most part by the connotations or emotionality of the words rather than by their characteristics of *Hi-* or *Lo-fit*.

TABLE III

MEAN DEVIATION FOR VERBS AND NOUNS ON ACTION VS. NON-ACTION ORIENTED SEMANTIC DIFFERENTIAL FACTORS

	Non-action		Action	
	Evaluation	Novelty	Potency	Activity
Verbs*	1.09	1.12	1.54	1.65
Nouns†	1.25	1.37	1.00	1.14

* wed, vote, steal, sin, reach. † bride, route, rust, root, barn.

Parenthetically, it is of interest to note that the distribution of nouns and verbs in the *Hi-* and *Lo-F* categories accords well with a suggestion made by Jespersen that words denoting movement would be one of the types very likely to manifest phonetic symbolism.[9]

Discussion. The present findings have several implications which bear attention. First, it was demonstrated once again, in the preliminary phases of the study, that under appropriate conditions it is possible to elicit consistent judgments of sound-meaning fit. The need for appropriate conditions, *i.e.* for a meaningful context in which such judgments are to be made, has been discussed elsewhere and will not be detailed here.[10] It is sufficient for our purpose to say that the meanings of the words themselves in the present study constitute the 'context' in which the judgment of whether a sound is 'appropriate' can be made. That is, one must know to what the sound is supposed to be appropriate before a meaningful judgment can be made. There is, however, a link between what can be said about these preliminary judgments and the interpretation of the semantic differential ratings which is of direct concern and will, therefore, be discussed more fully. This link derives from the likelihood that in both cases it is the word meaning which is precedent or gives rise to the observed judgment. Thus, in the case of the semantic differential ratings, we find that perceived meanings of words are not consistently influenced by 'fit' in any way which would suggest that phonetic symbolism is, in any sense, an independent influence. In other words, from our results, phonetic symbolism cannot be seen as having in any sense an existence prior to, or independent of, meaning categories, but must instead be considered a psychological response, initiated or suggested by (word) meanings, which then allows such meanings to become attached to otherwise meaningless sounds. Some evidence for this may be drawn from observations of such language phenomena as poetry, in which the evanescent appropriateness of sound is highly dependent upon the total context in which it occurs. Transferred to a less suggestive surround, those words whose sounds seem perfect to a given line would undoubtedly lose their sense of 'fit' and embeddedness in the total image. Even simply scrambling the words in a line of poetry would have the same effect. Phonetic symbolism, then, is highly dependent upon suggestion, a fact which was very early noted by Bentley and Varon.[11]

---

[9] Jespersen, *op. cit.*, 397.

[10] J. H. Weiss, The role of stimulus meaningfulness in phonetic symbolism, *J. gen. Psychol.*, in press.

The second major implication of our findings relates to Brown's functional hypothesis. While our study did not test that hypothesis directly, the suggestion of our findings is contrary to Brown's notion. Thus, it seems that words are not inherently high or low in phonetic symbolism until the suggestion is made, or a context implies, that such a sound-sense relationship is present. It would, therefore, be best to regard with caution any statement that words which are only *post hoc* high in phonetic symbolism have any spontaneous effect on the retention of those words or, for that matter, on other cognitive processes. In fact, several investigators who have attempted to discover such effects of phonetic symbolism, *e.g.* on the ease of learning 'appropriate' names to various objects,[12] on the recognition of figures to which variously symbolic phonetic sequences had been assigned as names,[13] or on the selection of names for various nonsense-figures have failed to do so.[14] Similarly, we fail to detect any consistent effect on the intensity of perceived meaning.

---

[11] Madison Bentley and Edith J. Varon, An accessory study of phonetic symbolism, this JOURNAL, 45, 1933, 76-86.

[12] A. E. Horowitz, The effects of variation in linguistic structure on the learning of miniature linguistic systems, Unpublished Doctoral dissertation. Harvard University, 1955.

[13] Roger Brown, *Words and Things*, 1958, 136.

It is recognized, of course, that the foregoing statement about Brown's hypothesis is based upon an indirectly related set of findings. It would be appropriate, therefore, to conclude this paper with a suggestion for a more direct test of Brown's theory. A paired-associates learning task in which nonsense-word pairs of varied 'appropriateness' were attached as names to objects varied on several physical dimensions or to pairs of polar adjectives which would serve as 'meanings' for the supposed 'words,' would constitute the basic procedure. From Horowitz's findings we would expect no differences in learning to mastery as a result of varying the 'fit' of pairs of the nonsense word.[15] One would predict, however, from Brown's position that recall would be aided by 'fit.' Thus, measures of retention over time would constitute the data to test the prediction.

SUMMARY

An attempt was made to assess the contribution, if any, of phonetic symbolism to the perceived intensity of word-meanings. Forty-five *S*s rated groups of words which were matched for 'emotionality' but varied on sound-meaning 'fit,' on Semantic Differential scales representing five factors.

On two factors, Evaluation and Novelty, words low in 'fit' elicited more extreme ratings, contrary to expectations based upon a report by Wertheimer who had demonstrated a correlation between high intensity of meaning and high 'fit.' On two other factors, Potency and Activity, high 'fit' words elicited more extreme ratings but this was traced to the fact that more high 'fit' words were verbs. On the basis of these findings, the hypothesis that phonetic symbolism adds to the perceived intensity of word meanings was rejected. Instead, it was argued that phonetic symbolism is dependent upon suggestion embodied in word-meanings. Results were discussed also in terms of their implications for Brown's functional hypothesis concerning the role of phonetic symbolism in the evolution of language and a more direct test of that hypothesis was suggested.

---

[14] D. K. Usnadze, Ein experimenteller Beitrag zum Problem der psychologischen Grundlage der Namengebung, *Psychol. Forsch.*, 5, 1924, 25-43.

[15] Horowitz, *op. cit.*

DOMINANCE OF THE TEMPORAL VISUAL FIELDS AT A SHORT DURATION OF STIMULATION

By HERBERT F. CROVITZ, VA Hospital, Durham, North Carolina,
and D. BARRY LIPSCOMB, Duke University

The specific topic of the present paper is binocular color rivalry. The paper reports alternatives to both rivalry and fusion which occur with a short duration of stimulation, the major alternative being a percept corresponding to the colors presented in the temporal visual fields.

The general question which forms the background of this study is the basic unsolved problem of how the binocular field is related to the monocular fields which must underlie it. The theory usually taught in psychology courses is the Fusion Theory. It assumes that there are corresponding areas of the two eyes, that fibers from each pair of such corresponding points meet in the occipital lobe and fusion occurs. Further, an isodynamic cell is assumed, one which can be fired by either or both fibers of each pair, and one which is responsible for sensation. This form of the Fusion Theory may be attributed to Cajal.[1]

There are three rivals to the Fusion Theory. First, there is the Suppression (or Replacement) Theory. It assumes that there are corresponding areas of the two eyes, that the two fibers from each pair of such corresponding points meet in the occipital lobe, but only one fiber of each pair is responsible for sensation—the sensation associated with stimulation of one of the pair of corresponding points is suppressed. One form of this theory has recently been espoused by Asher; another form of it has continually been put forward over the last half a century by Verhoeff.[2]

In opposition to both the Fusion Theory and the Suppression Theory is a Non-Physiological Theory. Its major champions have been Helmholtz[3] and Sherrington.[4] Both recognize the notion of corresponding areas of the two eyes and the convergence of afferent pathways, but they deny that such an arrangement has major relevance for sensation. Both assert that processes associated with sensation remain independent between the two eyes, and the binocular percept represents a psy-

---

* Received for publication November 12, 1962. The present work grew out of discussions of the senior author with Professors Karl Zener and Irving T. Diamond. Appreciation is expressed for their substantive suggestions.

[1] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 226-228; C. S. Sherrington, *The Integrative Action of the Nervous System*, 1961, 375-376.

[2] Harry Asher, *The Seeing Eye*, 1961, 209-230; F. H. Verhoeff, A new theory of binocular vision, *Arch. Ophthal.*, 13, 1935, 151-175; Panum's areas and some other prevailing misconceptions concerning binocular vision, *Transact. Amer. ophthal. Soc.*, 57, 1959, 37-47.

[3] Hermann von Helmholtz, *Treatise on Physiological Optics*, 3rd ed., in English trans., J. P. C. Southall (ed.), 3, 1925, 499.

[4] Sherrington, *op. cit.*, 377.

chological product of independently acting and independently elaborated monocular percepts. Like Sherrington,[5] Walls argues that the value of hemidecussation in the visual system of man is giving "control of the both sets of ocular muscles to single cortical loci, in the interests of smooth operation of the conjugated fixational and other . . . reflexes mediated via the visual cortex."[6] Thus, this theory places the relevance of anatomic facts not on seeing, but on moving the eyes; it leaves the problem of the relation of the binocular to the monocular fields unsolved.

In opposition to the above-mentioned theories stands one suggested by Edridge-Green.[7] In a pair of notes concerning binocular depth perception, he put forth the suggestion that when both eyes are open, the left eye predominates for the left visual field and the right eye predominates for the right visual field. In physiological terms, this would mean that the crossed fiber-system, mediating seeing the temporal visual fields, is dominant.

When this notion came to attention, puzzling data were at hand which it could explain. These data were collected in a replication of an experiment reported by Crovitz and Daves.[8] The stimuli consisted of a line of six numerals, three to the left of fixation and three to the right of fixation. They were exposed for 100 m.sec. in a Gebrands mirror tachistoscope. For a third of the trials, viewing was binocular; for a third, the right eye was occluded; for a third, the left eye was occluded. Twelve new *Ss* were added to the 14 tested originally. For the larger sample of 26 *Ss*, it was found that accuracy was significantly different in the right and left fields and among the three viewing conditions. The interaction between fields and viewing conditions was, furthermore, significant. Fig. 1 shows the percentage of numerals reported correctly under each viewing condition.

The finding relevant to the topic of the present paper is that when viewing is binocular, accuracy is just as great in the left field as it is when viewing with the left eye alone and accuracy in the right field is just as great as it is when viewing with the right eye alone. One possible interpretation of this finding is that when both eyes are open, the left eye 'sees' the left side and the right eye 'sees' the right side of the field of vision.

Were this theory indeed correct, a simple observation could confirm it. If a red card were presented to the left eye and a green card were presented to a corresponding area of the right eye, one should see a bipartite color field, red on the left and green on the right. this implictaion, however, is absurd. These conditions of viewing are well known to yield binocular color rivalry—a full field of red alternates over time with a full field of green. Or color fusion may occur, giving a full field of yellow or gray.

Yet, such a theory is consistent with the data shown in Fig. 1—data collected at a short duration of exposure. At 100 m.sec. voluntary eye-movement does not occur, and if Walls is correct, it is to eye-movements that one must look to assess the significance of the anatomic arrangement of the human hemidecussating visual system. We knew of no studies of binocular color rivalry at a short duration of exposure, arranged for such viewing, and were pleasantly astonished to experience the bipartite color field which was implied by the notion that at a duration of exposure shorter than the latency of voluntary eye-movement perception corresponds to those stimuli appearing in the temporal fields and leading to excitation of nasal retina and the crossed fiber system.

---

[5] Sherrington, *op. cit.*, 383.

[6] G. L. Walls, The lateral geniculate nucleus and visual histophysiology, *Univ. Calif. Publs. Physiol.*, 9, 1953, 42.

[7] F. W. Edridge-Green, Binocular vision, *J. Physiol.*, 48, 1914, xxxviii-xxxix, The stereoscopic appearance of certain pictures, *Nature*, 106, 1920, 375.

[8] H. F. Crovitz and Walter Daves, Tendencies to eye movement and perceptual accuracy, *J. exper. Psychol.*, 63, 1962, 495-498.

The present experiment is a systematic investigation of this discovery.

Method: (1) Apparatus. A specially constructed mirror tachistoscope was used for presenting the stimuli. It had three fields: a binocular fixation field; a field presenting stimuli to the left eye; and a field presenting stimuli to the right eye. Each field was 26 in. from the eyes. Illumination was given by cold fluorescent bulbs controlled by Scientific Prototype lamp driver and timer units. This apparatus allowed for alternate presentation of a fixation-point to both eyes (to control focus, convergence, and the line of regard) and independent stimuli presented simultaneously to the left and the right eye.

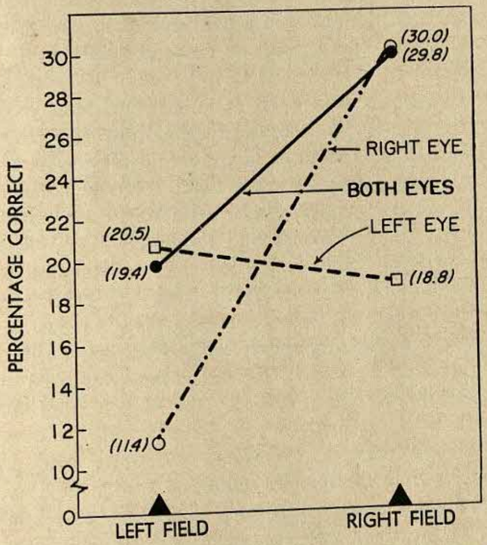


FIG. 1. PERCENTAGE OF NUMERALS CORRECTLY REPORTED UNDER EACH VIEWING CONDITION

(2) Stimulus-materials. There were two main stimulus-conditions, both designed to present different colors to corresponding areas of the two eyes. In the first, a full field of green (Munsell 5G5/8) was presented the left eye while a full field of red (Munsell 5R5/8) was presented to the right eye. On half the trials in this condition the colors were reversed so that a full field of red was presented to the left eye and a full field of green was presented to the right eye. Here, a perception of a split field would be expected, with the color presented to

the left eye appearing to the left of the color presented to the right eye. In the second condition, bipartite fields were shown to the two eyes. For half the trials the left eye was presented with a field with the left half green and the right half red while the right eye was presented with a field with the left half red and the right half green. On the other half of the trials the left eye was presented with a field with the left half red and the right half green while the right eye was presented with a field with the left half green and the right half red. Here, predominance of the crossed fiber system would lead to the perception of a full field of the color shown in the temporal field of each eye.

The binocular fixation-field consisted of a small black cross centered on a white field. All fields were rectangular and subtended 13° of visual angle horizontally and 10° of visual angle vertically, symmetric about the foveal fixation-cross.

(3) Procedure. A 'Ready' signal was given before each trial. The colored stimuli replaced the fixation-field for a constant exposure-duration of 100 m.sec. The stimuli were presented in a fixed random order with the restriction that no particular stimulus-arrangement was given more than twice in succession. In total, there were 80 trials, 20 presentations of each stimulus-arrangement. Except for the illuminated field appearing in front of him, *S*'s field of view was darkened.

The *S*s were told they would be shown some colors and that their task would be to report what they saw, with particular attention to the spatial arrangement of the perceived colors. They were also told they might see a field all red or all green; a split field with one color on the left of the other color; a series of stripes alternating in color; or one color apparently framed by another color. They were told further that they might see arrangements of color not included in these categories, and that in every case they should report what colors they saw in whatever spatial arrangements they appeared. The categories mentioned to the *S*s represented the vast majority of reports given by the *S*s in a pilot study to whom no structured instructions had been given.

(4) Subjects. Sixteen students, 13 women and three men, served as the *S*s. They were volunteers from the summer session of introductory psychology at Duke University in 1962. Prior to the experiment, visual screening-tests were igven. No person took part in the study who showed marked phoria.

Results. The stimuli were designed to present conflicting colors to corresponding areas of the visual system in order to explore the possible prevalence in vision of the color or colors leading to stimulation of the crossed fiber system. The results give clear evidence that such prevalence of stimulation of the crossed fiber system occurs.

In both the full-field and the split-field conditions, the most common report corresponded to the color presented in the temporal visual fields, which activates nasal retina and the crossed fiber system. In the full-field conditions the most common report was of a single field consisting of two colors, the color stimulating the left eye appearing to the left of the color stimulating the right eye. This was the result for 10 *S*s. In the split-field condition, the most common report was of a field consisting of one color—the color stimulating the nasal portions of the two eyes. This was the result for 8 *S*s. In most cases, under this condition, a thin line of the other color appeared vertically down the center of the perceived field. This perhaps occurred when the monocular fields were not in perfect alignment. This seems to be the most probable explanation, since the eyes tend to diverge when the stimuli are shown and the fixation-point is momentarily extinguished.

A second resolution of the conflicting stimulation corresponds to binocular rivalry. In the full-field condition all of one color was seen; and in the split-field condition, the field was split. A third resolution was the report of one color surrounding or framing the other color. A fourth resolution was reports of a series of vertical stripes of alternating colors. The latter occurred for only a few *S*s but made up 80% of the 'other' category. These data are shown in Table I.

TABLE I

COLORS SEEN WITH FULL-FIELD AND SPLIT-FIELD STIMULATION

	Full-field					Split-field				
S	One eye	Temporal field	Nasal field	Frame	Other	One eye	Temporal field	Nasal field	Frame	Other
1	0	19	15	6	0	38	0	0	0	2
2	4	32	1	1	2	20	12	2	4	2
3	9	26	3	1	1	22	11	0	5	2
4	0	0	0	0	40	1	0	1	0	38
5	28	6	0	0	6	6	20	11	0	3
6	3	19	1	11	6	4	24	1	6	5
7	3	19	3	12	3	30	4	3	0	3
8	0	30	2	0	8	35	2	0	0	3
9	6	22	1	1	10	2	27	0	0	11
10	0	12	0	6	22	0	24	1	12	3
11	0	9	2	29	0	28	0	0	11	1
12	0	33	1	5	1	1	26	1	10	2
13	0	14	1	22	3	6	24	3	0	7
14	0	32	2	3	3	9	29	1	0	1
15	4	10	12	0	14	4	16	15	0	5
16	0	17	5	14	4	1	13	0	1	25
Total	57	300	49	111	123	207	232	39	49	113
%	8.9	46.9	7.7	17.3	19.2	32.3	36.2	6.1	7.7	17.7

Sometimes the perception corresponded to the stimuli in the nasal rather than the temporal field. When the report corresponded to input from either the crossed or uncrossed systems, correspondence to the crossed fiber system occurred more frequently than correspondence to the uncrossed fiber system. In the full-field condition, $t = 5.66$, $p < 0.001$; in the split field condition, $t = 4.32$, $p < 0.001$. In each condition this relationship held for 14 of the 16 *S*s.

Reports corresponding to stimulation of one eye (rivalry) occurred more frequently in the split-field than the full-field condition for 12 of the 16 *S*s ($t = 2.35$, $p < 0.05$). Binocular fusion never occurred; *i.e.* there was no report of a full field of yellow, gray, or white.

Discussion. The results of the present study reveal alternatives to rivalry at a short duration of stimulation. This finding has immediate relevance to any theory of binocular rivalry and its eventual importance may be in development of a correct understanding of the relation of the binocular field to the monocular fields which underlie it.

Woodworth discussed the type of experiment in which discrepent stimulation is applied to corresponding parts of the two retinas. He asked what response will the brain make to such a conflict of cues, and outlined the alternatives; namely, it might disregard one retina and respond only to the other (rivalry), or it might respond to the combination in several other ways: by fusion of the monocular fields, by seeing one through the other, or, where possible, by getting a depth effect.[9] The present data reveal yet other alternatives when stimuli are presented for 100 m.sec. These alternatives appear to have direct physiological meaning. The most frequently occurring one corresponds to stimuli exciting nasal retina, or less commonly, temporal retina.

An appealing possibility with respect to rivalry is suggested by the less commonly occurring percepts: one color framing another and a series of stripes. A working hypothesis might be that these indicate stages in the temporal development of the binocular field. For example, an initial percept might be determined by the distinction between the crossed and uncrossed afferent fiber systems, that later in time there is partial spatial intrusion of the field of one eye upon the other, and that later still rivalry or fusion occurs. Whether the characteristics of the binocular field develop over time, and whether this represents a correct sequence of changes, remains a task for empirical determination.

After the present experiment was completed, we found a paper which reported a similar set of observations.[10] Köllner noticed that when holding a red glass before one eye and a blue glass before the other, the initial percept on opening his eyes was a bipartite color field, the color before the left eye appearing to the left of the color before the right eye. In his observations, time of exposure to the colors was not controlled, but he noted that the color halving of the visual field remains in existence only a short time, quickly giving way to rivalry.

Köllner argued that this phenomenon implies a functional superiority of the temporal visual fields, and our data likewise point to a dominance of the crossed fiber visual system, at least at some early stage in the processing of visual information. There is scattered evidence of such superiority in the physiological literature.[11] The present data invite more specific exploration of physiological variables involved in the nasal-temporal distinction among human *Ss*.

---

[9] R. S. Woodworth, *Experimental Psychology*, 1938, 572-573.

[10] H. Köllner, Das funktionelle Überwiegen der nasalen Netzhauthälften in gemeinschaftlichen Sehfeld, *Archiv. f. Augenhk.*, 76, 1914, 153-164.

[11] A good review is given by Walls, *op. cit.*, 88-93. For some recent work, see B. D. Burns, Woodburn Heron, and Bernice Grafstein. Responses of cerebral cortex to diffuse monocular and binocular stimulation, *Amer. J. Physiol.*, 198, 1960, 200-204; R. W. Doty, Potentials evoked in cat cerebral cortex by diffuse and punctate photic stimuli, *J. Neurophysiol.*, 21, 1958, 437-464; D. H. Hubel and T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.* 160, 1962, 106-154.

One limitation of the design of the present experiment is that we did not measure the relative size of the apparent color-extents. A report of "red to the left of green" could reflect a field of half red and half green and it could also reflect a field with a mere border of red to the left of a large area of green. The empirical question which remains is the extent to which nasal-field colors might intrude even at 100 m.sec.

A problem raised by the present study concerns the processes underlying binocular rivalry. The present data point to the possibility that at short exposure-times, perception may be simply related to anatomic distinctions. A percept corresponding to stimulation of temporal fields and the crossed fiber system, for example, seems reasonable to the extent that the anatomy on which such a distinction might be based is known to exist. But a percept corresponding to the stimulation of one eye, as is common in the rivalry occurring at long exposure-times, does not appear reasonably related to the anatomy of the afferent visual paths, since 'an eye' loses its unity as fibers leading from it take diverging paths to cortical loci on opposite cerebral cortices.

One speculation might place the emergence of rivalry as the consequence of an innervation for conjugate eye-movement. The present data show that, at an exposure-time faster than the latency of voluntary eye-movement, alternatives to rivalry occur. Sherrington and Walls argue that the anatomy of the visual system in man appears designed to facilitate efferent messages controlling eye-movement. It would not be strange if eye-movements were, then, related to rivalry. Data supporting this view does not, however, exist. Furthermore, Peckham presented data which led him to deny a relation of eye-movement to shifts in rivalry.[12] Yet, Köllner noted that when looking through colored glasses and waiting until binocular color-fusion occurs, a lateral conjugate eye-movement led to seeing only the color before one eye. He suggested that insignificant unintentional lateral eye-movements may assist the color of one eye to dominance. Further investigation of the relation of eye-movements, or of tendencies to eye-movements, to characteristics of binocular vision are indicated.

SUMMARY

Binocular color rivalry was studied with 100 m.sec. exposures. Among the percepts reported were the colors presented to one eye, the colors presented to the nasal or the temporal fields, one color framed by another color, and a series of stripes of alternating colors. The most common percept corresponded to the colors stimulating nasal retinas. The data suggest the value of considering that the binocular field develops over time from the monocular fields with the crossed fiber system dominant at an early stage.

---

[12] R. H. Peckham, Eye-movements during 'retinal rivalry,' this JOURNAL, 48. 1936, 43-63.

'SPREAD OF EFFECT' WITHOUT GUESSING SEQUENCES

By ELIZABETH ALFERT, University of California, Berkeley

The 'spread of effect' refers to a gradient of repetitions of errors following and preceding a reinforced response. Thorndike interpreted this phenomenon as evidence for the automatic influence of reward.[1] Later studies have usually confirmed the existence of an after-gradient whereas a significant fore-gradient has been found only in a few instances.[2] The most widely accepted explanation of the after-gradient rejects Thorndike's original hypothesis and attributes the 'spread of effect' to the nonrandom character of *Ss*' guessing sequences. Repetition of the rewarded response on a test trial reinstates the sequential response-biases which had been present on the training trial.

The guessing-sequence explanation was first proposed by Jenkins and Sheffield,[3] and later supported by Jenkins and Cunningham, Sheffield, Fagan and North, and Sheffield and Jenkins.[4] While the operation of response-biases was clearly demonstrated, it does not follow that the observed gradients are entirely determined by guessing sequences. A recent experiment by Postman has cast doubt on the hypothesis of 'guessing-sequence' as an exhaustive explanation of the 'spread of effect.' According to this hypothesis repetition of the rewarded response is a necessary condition for the appearance of an after-gradient. Postman obtained a significant after-gradient in a situation in which repetition of the reinforced responses was greatly reduced by introduction of a delay between training and test.[5] In neither the immediate nor the delayed tests were the after-gradients dependent on repetition of the rewarded response.

The purpose of the present investigation was (1) to provide a further

---

* Received for publication October 11, 1961. The author is indebted to Dr. Leo Postman for his guidance and criticism throughout the investigation.

[1] E. L. Thorndike, A proof of the law of effect, *Science,* 77, 1933, 173-175; An experimental study of rewards, *Teachers Coll. Contr. Educ.,* 1933 (No. 580), 1-72.

[2] M. H. Marx, Spread of effect: A critical review, *Genet. Psychol. Monogr.,* 53, 1956, 119-186.

[3] W. O. Jenkins and F. D. Sheffield, Rehearsal and guessing habits as sources of the spread of effect, *J. exp. Psychol.,* 36, 1946, 316-330.

[4] W. O. Jenkins and L. M. Cunningham, The guessing sequence hypothesis, the 'spread of effect' and number guessing habits, *J. exp. Psychol.,* 39, 1949, 158-168; F. D. Sheffield, 'Spread of effect' without reward or learning, *J. exp. Psychol.,* 39, 1949, 575-579; C. A. Fagan and A. J. North, A verification of the guessing sequence hypothesis about spread of effect. *J. exp. Psychol.,* 41, 1951, 349-351; Sheffield and Jenkins, Level of repetition in the spread of effect, *J. exp. Psychol.,* 44, 1952, 101-107.

[5] Leo Postman, Spread of effect as a function of time and intraserial similarity, this JOURNAL, 74, 1961, 493-505.

test of the guessing-sequence hypothesis and (2) to determine the specific stimulus-conditions which are necessary for the occurrence of the phenomenon. Specifically, the present experiment investigates the effect of changing the stimulus-word in the rewarded position between training and test while leaving all the other stimulus-words intact and in their previous positions. This procedure leaves the general serial context unchanged while reducing repetitions of the rewarded response and the guessing sequence anchored to it.

Methods. (1) Learning material. The stimulus-items were 21 words of low intralist-similarity, as described by Duncan.[6] There were four different random orders which were used equally often. The words were presented on a memory-drum at a 5-sec. rate. The length of the exposure-interval was 2.5 sec. There was a 12-sec. interval between the two presentations. The *S*s were told that they were participating in an experiment on learning to associate numbers with words. For purposes of habituation to the procedure there was a short period of practice during which typewritten symbols were shown on the memory-drum, to which *S*s responded with random letters of the alphabet. The *S*s were then instructed to respond to each word with a number between 2 and 9. They were to choose these numbers at random, make guesses the first time they saw the words, and would be told 'Right' or 'Wrong' after each response. The *S*s were told that the series of words would then be shown to them again and were asked to repeat the numbers that had been correct, while changing the numbers that had been wrong. There were one training trial and one test-trial.

(2) Experimental design. To determine the conditions essential for producing the 'spread of effect' an attempt was made systematically to eliminate possible sources of the effect. There were four experimental conditions. The training trial was the same for all conditions whereas the procedure on the test-trial was varied. The fifth, eleventh, and seventeenth responses were rewarded during the training trial by *E*'s announcement of 'Right,' while all other responses were called 'Wrong.' With Group A the conventional procedure for measuring the 'spread of effect' was used. The series used in the training trial was repeated in the test-trial and *S*'s responses were rewarded at the same positions regardless of what they were. Group B was also given the same series on the training trial and test-trial but *S* was not rewarded for any of his responses, E saying 'Wrong' to all his responses during the test. For Groups C and D new words were substituted in the test in the fifth, 11th and 17th positions, so that the stimulus-words for the rewarded responses were not repeated during the test while the general context of the list remained the same. For Group C responses in these critical positions were rewarded while for Group D an announcement of 'Wrong' was made by *E* after all of *S*'s responses. Thus for Groups C and D the change in stimuli reduced the probability of repetition of the rewarded responses and hence of guessing sequences anchored to these responses. For Group D the rewards during the test-trial were eliminated as well

---

[6] C. P. Duncan, Stimulus generalization and spread of effect, this JOURNAL, 64, 1951, 585-590.

so that a guessing sequence could neither be anchored to the rewarded response nor to *E*'s announcement of 'Right'.

(3) Subjects. There were 96 *S*s in the experiment, 24 in each experimental condition. The *S*s were students at the University of California taking courses in psychology which required participation in experiments.

Results. Table I shows the number of repetitions of the rewarded response as well as the repetitions preceding it in Positions −1 and −2 and the repetitions following it in Positions +1 and +2. In Conditions A and B the list of words remained unchanged during training and test, and repetition of the rewarded response was favored. In Conditions C and D the stimulus for the rewarded response was changed and, as a consequence, repetition of the previously rewarded response was no more frequent than repetition in any other position. (In Condition D there were slightly fewer repetitions than in Position −1 or −2, in Condition C slightly more). Thus recall of the position of the rewarded response did not lead to its repetition when the stimulus was absent.

TABLE I

FREQUENCY OF REPETITIONS

Position	Group A	Group B	Group C	Group D
−2	6	7	9	10
−1	12	8	10	10
R	28	38	14	8
+1	16	14	15	16
+2	10	7	10	11

No significant fore-gradient was found under any of the conditions, though Condition A shows a gradient in the expected direction.

Comparison of the frequency of repetitions in Positions +1 and +2 shows a declining gradient under all conditions. An analysis of variance yields an *F*-ratio of 4.95 ($df = 1/92$) for Positions, which is significant at the 5% level, while the *F*-ratios for Conditions and the interaction, Conditions × Positions, are not significant. An after-gradient following reward is, therefore, demonstrated which does not vary significantly from condition to condition. Neither substitution of new stimulus-words for the rewarded responses nor a reversal in the reinforcement for the previously rewarded response, nor the combination of these changes eliminated the 'spread of effect.'

The frequency of repetitions in Position +1 was also compared against an empirical baseline determined by the mean frequency of repetitions in Positions −1 and −2. An analysis of variance yields an *F*-ratio of 5.45, ($df = 1/92$) for Positions, significant at the 5% level, while *F*-ratios for Conditions and Conditions × Positions are not significant. Since the interaction is not significant, repetitions in Position +1 are more frequent than in Positions −1 and −2, regardless of whether the repetitions of the previously rewarded responses do or do not significantly exceed that baseline.

Conditions D and A represent the largest difference in the method of testing after identical training trials. Figure 1 presents the results for these two conditions and shows the reduction of repetitions of the previously rewarded response when a new stimulus is substituted and called 'wrong by the experimenter and the increase of repetitions in Position +1. The magnitude of the after-gradient is very nearly the same in both Conditions.

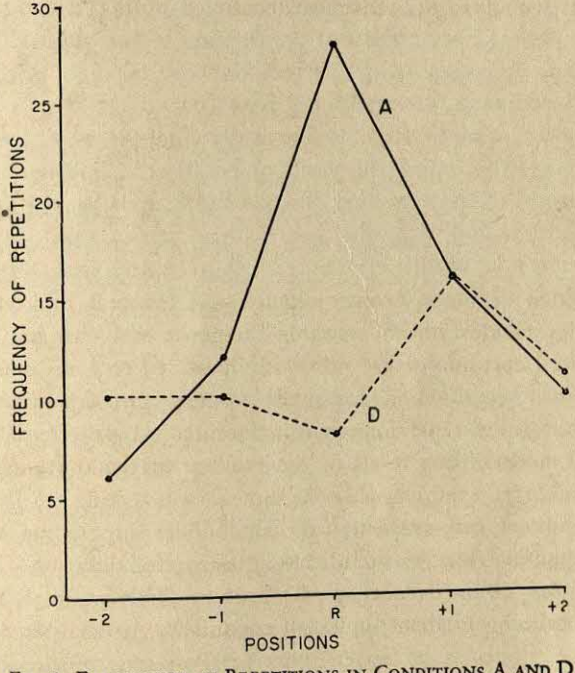


FIG. 1. FREQUENCY OF REPETITIONS IN CONDITIONS A AND D
(Stimuli and rewards were repeated in Condition A. New stimuli were substituted and rewards were eliminated in Condition D.)

Discussion. Marx in his review of the 'spread of effect' has discussed the major alternative theoretical interpretations of the Thorndike effect.[7] The guessing-sequence hypothesis and the hypothesis of serial response-

---

[7] Marx, *op. cit.*, 119.

response reinforcement are the main theories opposing Thorndike's explanation of the spread of effect in terms of the automatic strengthening of S–R connections. Both these theories assume that the after-gradient is contingent on the repetition of the rewarded response. This experiment, however, has shown that the occurrence of the phenomenon does not depend on the repetition of either the stimulus or the response, or the reward itself, in the previously rewarded position. This finding indicates that only the appearance of the stimulus in Position +1 is necessary to elicit the repetition. The S–R connection following the 'Right' must have been strengthened during the first presentation of the words. This result is in perfect agreement with Thorndike's view, though there was no reliable evidence for a fore-gradient.

The present findings seem to be contrary to those of Jenkins and Cunningham,[8] who varied the positions of all stimuli between training and test except in the rewarded position and found that the appearance of a 'spread of effect' depended on repetition of the rewarded response. This difference in results may have various explanations. (1) Jenkins and Cunningham used a number of successive test-trials which may have obscured the effect. (2) By changing the positions of all stimuli except those in the rewarded positions they drastically changed the serial context of individual items. (3) It is questionable whether rewards in Thorndike's sense were used in this experiment since the reinforcement consisted merely of an instruction to repeat a given response to a stimulus on the next trial.

It might be argued that reward in Thorndike's sense is necessary since it brings about a temporary increase in drive following the announcement of 'Right' which may facilitate learning. This would be in accord with Spence's theory that emotional responsivity, as expressed by autonomic activity, can act as an energizer.[9] Such responsivity would be increased by greater isolation of rewards which would account for Zirkle's results on the effects of isolation,[10] and would also explain Jenkins and Postman's finding that isolation without a reward did not produce a 'spread of effect.'[11] An interpretation of the 'spread of effect' in terms of drive would account for the asymmetry of the effect since an increase in drive would occur only following the reward.

---

[8] Jenkins and Cunningham, *op. cit.*, 158.

[9] K. W. Spence, A theory of emotionally based drive (D) and its relation to performance in simple learning situations, *Amer. Psychologist*, 13, 1958, 131-141.

[10] G. A. Zirkle, Success and failure in serial learning. 11. Isolation and the Thorndike effect, *J. exp. Psychol.*, 36, 1946, 302-312.

[11] Jenkins and Postman, Isolation and 'spread of effect' in serial learning, this JOURNAL, 61, 1948, 214-221.

An alternative hypothesis would attribute spread to a failure of discrimination. Uncertainty about what was right or wrong could induce repetition of responses close to the rewarded position. This would mean that better learners should show less of a 'spread of effect.' The number of *S*s in Condition A is too small to permit a reliable check of this interpretation but the fact that the *S*s who repeated more of the rewarded responses also had more repetitions in Position +1 gives no support to this hypothesis with regard to th eafter-gradient. There were, however, more repetitions in Position −1 among Ss who did not learn the rewarded responses, and it seems possible that whenever a fore-gradient has been found, it was the result of uncertainty of discrimination. This may have been the case in Duncan's experiment in which lists of word of low and high similarity were used. Duncan found a significant fore-gradient as measured by number of different responses in a series of test-trials for the list of high-similarity words.[12] In this list repetition in Position −1 was more frequent when the rewarded response was not repeated.

SUMMARY

This experiment investigated the conditions necessary to produce a 'spread of effect.' Subjects were asked to associate numbers to a series of 21 words. Three responses were reinforced during training. The conditions of the test-trial were varied systematically. A significant gradient of repetitions of errors occurred in the test-trial even when the rewarded stimulus, the rewarded response and the reward were not repeated. The frequency of repetitions in the position following reward was significantly above an empirical baseline even when that of the previously rewarded response was not.

---

[12] Duncan, *op. cit.,* 585.

HERING'S ILLUSION WITH IMPOVERISHMENT OF THE STIMULUS IN SCOTOPIC AND PHOTOPIC VISION

By A. J. MARSHALL and VINCENT DI LOLLO,
University of Western Australia

Theoretical accounts of the geometrical illusions differ in the emphasis placed on the over-all configuration as against the accumulation of effects from discrete parts of the stimulus. Orbison follows Gestalt tradition in accounting for the distortion of geometrical forms in terms of forces resulting from the patterning of figure- and background-lines.[1] The alternative position is illustrated by the classical explanation of line illusions in terms of overestimation of acute angles, and by more recent attempts to relate distortions of lines to sizes of angles at intersections.[2] Studies using microstimulation to establish deformations in fields of retinal induction have so far left the issue unsettled.[3] Related phenomenological studies of illusions have, however, stressed the critical importance of small parts of the figure, as in Kobayashi's finding that Poggendorff's illusion tended to disappear when the ends of the oblique lines meeting the verticals were bent slightly outward near the point of contact.[4]

One way of attempting to decide between the rival theories is to determine the extent to which illusions are affected by loss of detail through impoverishment of the stimulus. Some characteristic effects of impoverishment are common to both night- and day-vision, but there may be differences in the order of emergence of various properties of the stimulus as conditions of viewing improve in the two situations. If, for example, a regular checkerboard is distorted by blacking over one of its white squares, the distorted region so formed can be seen as a spot on an apparently homogeneous, unpatterned field when viewed with reduced angular size in day-vision. With larger angular size, but reduced lumi-

---

* Received for publication January 17, 1962. The authors are indebted to Dr. R. H. Day for help and criticism.

[1] W. D. Orbison, Shape as a function of the vector field, this JOURNAL, 52, 1939, 31-45.

[2] Anna Berliner and S. Berliner, The distortion of straight and curved lines in geometrical fields, this JOURNAL, 61, 1948, 153-166; E. B. Titchener, *Experimental Psychology*, 2, part 2, 1910, 315.

[3] J. W. Gebhard, Motokawa's studies on electrical excitation of the human eye, *Psychol. Bull.*, 50, 1953, 73-111.

[4] T. Oyama, Japanese studies of the so-called geometrical optical illusions, *Psychologia*, 3, 1960, 7-20.

nance, the same region cannot be detected until some degree of patterning of the checkerboard is also visible.[5] In general, the differences between night- and day-vision appear to depend on the greater tendency for night-vision to average out smaller detail and to align centers of density of regions of excitation, and for day-vision to make more use of contrast, to segregate regions and to form contours despite the small angular sizes involved.[6]

In the present investigation, the amount of illusion of a Hering-figure has been measured under the following two modes of impoverishment of the stimulus: (a) at various scotopic levels of luminance, from a constant viewing distance; and (b) from various distances with lighting held constant at a photopic level. On the assumption that small regions of the figure are of critical importance, it was predicted that the illusion would be very much reduced, and possibly absent, under certain conditions of impoverishment of the stimulus with both night- and day-vision, even though the background-lines could still be discriminated. Since, at scotopic levels, local detail emerges relatively slowly as compared with over-all patterning, it was also predicted that, while a certain limited range of acuities might well be critical for establishing the illusion in day-vision, its development in night-vision would proceed gradually over a wide range of degrees of impoverishment.

METHOD

(1) Apparatus: (a) Hering-figure. The figure consisted of two horizontal bars set on a background of lines fanning out from a point at the center of the figure. The background-lines, each 6.2 mm. in width, were printed in black ink on a white card 30.5 × 66.0 cm. Two horizontal brass bars, painted black and having a 6.2 × 6.2 mm. square section, were set closely against the card. The bars were fixed at the center of the apparatus, with 10.2 cm. separation of centers of bars. They could be bent symmetrically by bringing the ends closer together or by pressing them farther apart by means of a worm operated by *E* at the rear of the apparatus. A scale with a pointer indicated the distance separating the extremities of the bars. When the bars were set parallel, they made angles of approximately 20°, 40°, 60°, 90°, 60°, 40°, and 20° with the seven background-lines. The ends of the apparatus were screened by white cards 9.0 × 30.5 cm., framing a Hering-figure 49.0 × 30.5 cm. This figure, set in the fronto-parallel plane, was viewed by reflected light. In the distance-viewing condition, two 300-w. lamps, 1 m. apart and 30.0 cm. away from the figure, gave a fairly even lighting of approximately 160 ft.L. as measured by an SEI photometer. The screens at the ends of the figure threw a narrow

---

[5] R. S. Ives, Perceptual organization under different conditions of viewing, Unpublished thesis, University of Western Australia, 1961.

[6] R. H. Day, Application of the statistical theory to form perception, *Psychol. Rev.*, 63, 1956, 139-148.

shadow: no visible shadows were thrown by the bars themselves. The lamps were set in reflectors, 23.0 cm. in diameter, at approximately 135° from *S*'s line of sight, the inner edges of the reflectors being nearly in line with the outer edges of the two cardboard screens. The appaartus was mounted on a trolley 70.0-cm. high, which was covered with a black cloth. In the dark room two 60-w. lamps were set in the same reflectors, and in approximately the same position relative to the test-surface. The luminance was altered by means of a variac.

(b) Acuity-targets. Six 'E' visual targets were drawn in black ink on separate sheets of white, non-glossy cards 33.0 × 33.0 cm. These letters were of standard shape with the central bar shorter than the other two by the width of one bar. The widths of the bars were 8.0, 6.2, 4.4, 3.5, 2.6, and 1.7 mm. on the six different cards. Two further targets on white cards 33.0 × 20.0 cm. presented angles of 20° and 24° respectively. These angles were bounded by lines 6.2 mm. × 15.0 cm. printed in black ink. The lesser angle was equal to that of the extreme radial lines in the Hering-figure: an angle of 24° could be obtained by rotating an extreme radial line through its own width at the point of intersection with a horizontal bar.

(2) Subjects and procedure. The *Ss* were 24 university students who had at least 20/20 vision without correction, and no pronounced astigmatism. They were randomly allocated to two groups of 12 *Ss* each. The 12 experimental *Ss* viewed the Hering-figure in two different sessions of dark room and of distance-viewing, and told *E* how to adjust two horizontally placed bars to appear parallel against a background of lines. The 12 control *Ss* carried out the same procedure, except that they adjusted the bars without the background-lines. The amount of illusion for each *S* was measured at five levels of luminance in a dark room and at five viewing distances with photopic vision. Each level of luminance or distance used for an *S* was that at which he could just resolve one of the five largest 'E' targets. These sizes served to define degrees of impoverishment of the stimulus, since, whether the illusion was to be measured in night- or day-vision, those conditions at which *S* could just resolve a larger acuity-object were more difficult (with the target farther away or the level of luminance lower) than those at which a smaller acuity-object could be resolved. For each condition, *S* made 11 settings from 11 starting positions of the bars chosen randomly for each *S*. Of the 12 *Ss* in each group, 6, chosen at random, had the dark room session first and the distance-viewing second; the order was reversed for the 6 *Ss* remaining. Within each of these subgroups, the measurement of the illusion for 3 *Ss* began with the poorest viewing condition and proceeded to the most favorable, while for the other 3 *Ss* the order was reversed. Binocular vision with natural pupils was used throughout.

(a) Dark-viewing. During the first 30 min., while dark-adapting, *S* was given some practice at naming the orientation of the six 'E' targets, making six judgments with each. For those *Ss* who were to have a series of 'improving' conditions during the test proper, practice began with an 'E' of 1.7 mm. and ended with one of 8.0 mm. This sequence was reversed for the other *Ss*. At the end of this period, *E* determined for each *S* the setting of the variac (level of luminance) at which the first of the five largest 'E' targets could just be resolved. To do this the Hering-figure was covered with a white card 75.0 × 33.0 cm., and the appropriate target was held up in front of this in one of the four main orientations predetermined for each session by the use of a table of random numbers. The criterion adopted for

the resolution of the target was four out of four or seven out of eight correct judgments. The chosen level of luminance was that at which this target could be resolved but at which *S* was not successful with the next smaller. *S* then made five judgments of which was the greater of the two test-angles shown simultaneously, one above the other, in random order. The 11 settings of the horizontal bars were then made following these instructions:

> *Instructions.* You will see two horizontal lines ["as well as other lines" was added for the experimental group]. These can be changed from being parallel to bowing in or bulging out at the center. Your task is to direct me to set these lines parallel. When they appear to bulge out at the center [*E* illustrated with his hands] you should tell me to move them in at the center and tell me to stop when they look parallel. When they bow in, tell me to move them out and to stop when they appear to be parallel. You can always ask me to go back if I take them too far. If they look to be parallel when first shown to you, just tell me that they don't need any adjustment. I want you to shut your eyes after completing each setting.

These procedures were repeated at each of the remaining four levels of impoverishment of the stimulus. *S* was seated at a distance of 3 m. with the figure at eye-level when using the chin-rest. The session lasted approximately 1.5 hr.

(b) Distance-viewing. A corridor 122-m. long and 1.8-m. wide was used for the distance-viewing (photopic vision) sessions. Besides artificial light at the test-surface, the corridor had natural indirect lighting. *S* remained seated with a chin-rest at one end of the corridor and spoke to *E* through a telephone. The trolley carrying the apparatus was moved by *E* to appropriate distances determined for each *S* by testing with the 'E' targets. Instructions and procedures were closely comparable to those for dark viewing. The session lasted approximately 1¼ hr.

RESULTS

For each of the five degrees of impoverishment of the stimulus, in both dark and distance-viewing conditions, *E* recorded the 11 settings of the horizontal bars of the Hering-figure. These scores in millimeter separation of the ends of the bars were averaged to give a score for each *S* who thus had five scores for the session in the dark room and five scores for distance-viewing. Fig. 1. gives the average scores at each of the five degrees of impoverishment of the stimulus for the 12 experimental *S*s, and the corresponding average scores for the 12 control *S*s.

Table I is a summary of two analyses of variance carried out on the scores for experimental and control *S*s for dark- and distance-viewing separately. In dark-viewing there was a highly significant over-all amount of illusion for the experimental group. The highly significant effect of impoverishment of the stimulus 'E' on the magnitude of illusion ($p <$ 0.001) loses its relevance when compared with the significant C × E interaction. Taken together with the information supplied in Fig. 1, this interaction indicates that a significant amount of illusion was found only at some levels of impoverishment, and then only in the experimental group. Practically the same picture is presented by the scores obtained in distance-viewing. It must be noted, however, that, in this case, the over-all score indicates a significant amount of negative illusion.

The slope of the linear component of the performance of each *S* was computed separately for both dark- and distance-viewing. Each *S* thus had two slopes: one for the dark- and one for the distance-condition. Without exception, individual slopes were steeper in the case of distance-viewing. With 12 cases, the difference in slopes is highly significant $(p = 1/2^{12} = 0.0002)$. No incorrect judgments were made in discriminating the two test-angles at any level of impoverishment.

DISCUSSION

As shown in Fig. 1, the amount of normal illusion tended to be reduced with impoverishment of the stimulus-figure. With both scotopic and photopic vision, certain levels of impoverishment were reached at which *S*s perceived no illusion whatever. It could be argued that under such conditions, the radial lines were not discriminable, and that the situation was, therefore, virtually the same as for the controls. This objection is weakened, though not refuted, by the consideration that the bars of the Hering-figure and of the test-angles, although equal in width to the radial lines, were discriminated sufficiently to maintain a low variability of bar-settings and a high accuracy of angle-discrimination at all levels of impoverishment. It should be noted that this absence of illusion has been obtained by the method of impoverishing the whole stimulus. Elsewhere, it has been shown that illusory effects may be obtained when the inducing background-patterns are presented subliminally provided that the figure of the illusion is itself clearly perceived.[7]

In the dark room, the development of the illusion was gradual, and showed no tendency to depart from linearity over a range of conditions permitting a threefold increase in acuities (0.11 to 0.33). All the *S*s, including those seeing the figure for the first time at the lowest level of luminance, consistently reported seeing a figure which had the general characteristics of horizontal bars and of lines fanning out from a central region. Since this over-all patterning was discriminable throughout, the development of the illusion seems to be related to the change from indefinite to definite form which enabled the emergence of detail at the intersections of the radial lines with the bars. It would be possible to relate the stages of perception more directly to amount of illusion by requiring the *S*s to discriminate between the present Hering-figure and variants having radial lines with serrated edges and different sizes of detail. While not decisive, the results obtained with the present method tend to favor the explanation of the illusion in terms of discrimination of detail at intersections, since the rival position of over-all directional fields would demand the further explanation of why definite form of the bars is required for the adequate development of the illusion.

---

[7] G. J. W. Smith and M. Henriksson, The effect on an established percept of a perceptual process beyond awareness, *Acta Psychol.*, 11, 1955, 346-355; M. E. Goldstein, Subliminal perception with optical illusions, *J. gen. Psychol.*, 62, 1960, 89-101.

TABLE I

SUMMARY OF TWO ANALYSES OF VARIANCE

Source	*df*	Dark		Distance	
		MS	*F*	MS	*F*
Among the *S*s					
Condition (Exp-Cont.), (C)	1	991.00	24.81*	407.00	17.32*
Error (b)	22	39.95		23.50	
Within the *S*s					
Impoverishment (E)	4	140.75	50.01*	1409.50	92.55*
C×E	4	145.75	51.87*	1352.75	88.82*
Error (w)	88	2.81		15.23	
Total	119				

* $p<0.001$.

In distance-viewing, the development of the illusion was continuous over the range of conditions but the slope of the curve was steeper than the corresponding curve in the dark room. The interpretation of this finding is complicated by the presence of an illusion which is the reverse of the normal Hering-illusion, with the smaller angular sizes in distance-viewing. Trials with other figures suggested that this effect is not unique to the present arrangements of lines nor to the use of radial lines equal in width to the horizontal bars. Although the literature indicates that the normal Hering-illusion must be complexly determined, this effect was unexpected.[8] It is possible that a common principle of perceptual organization in photopic vision may be operating in the occurrence of both forms of the illusion. With reduced sizes of figure, there is a greater tendency toward segregation of regions rather than toward primary patterning, and at high levels of impoverishment the regions thus segregated are relatively large. Subjective reports indicated that the *S*s tended to perceive as one region all of the area bounded by the horizontal bars and the most slanted radial lines beyond the point of intersection. This could give a diverging effect to the horizontal bars which, therefore, would have to be set closer together at the extremities to appear parallel, thus giving rise to negative illusion. At closer distances, with less impoverishment of the stimulus, smaller regions of intersections of bars and lines could be segregated with the consequent emergence of the normal illusion. It is also possible that negative illusion may arise directly from physical blurring of the optical image itself at small angular sizes, and thus be a secondary phenomenon unrelated to normal Hering-illusion. This seems less likely when one considers the occurrence of a significant amount of negative illusion under conditions at which it was possible to resolve a Snellen 'E' having detail (4.4 mm.) smaller than that of the lines and bars of the Hering-figure. Furthermore, some trials using tachistoscopic presentation have shown that

---

[8] Berliner and Berliner, The distortion of straight and curved lines in geometrical fields, this JOURNAL, 61, 1948, 153-166; J. L. Zajac, Some investigations of the so-called "Geometrico-optical illusions," *Acta Psychol.*, 13, 1957, 140-150.

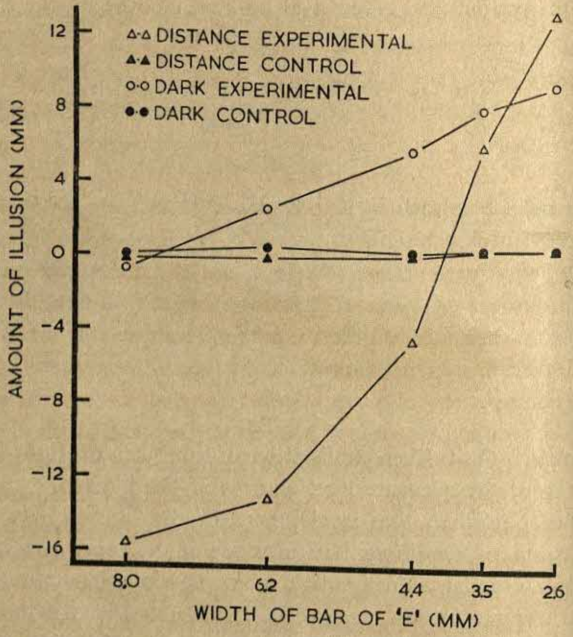


FIG. 1. THE AMOUNT OF ILLUSION IS THE AVERAGE SEPARATION OF THE BARS AT THE EDGE OF THE HERING-FIGURE, LESS THEIR SEPARATION WHEN PARALLEL

Each point represents the average of the results of 12 *Ss*. Larger widths of bar of 'E' correspond to higher degrees of impoverishment of the figure.

negative illusion can be obtained in place of normal positive illusion simply by reducing the time of exposure.

Any comparison between night- and day-vision, using the present method of equating degrees of impoverishment of the stimulus, must take into account the grossly different sizes of retinal area involved in the two situations. In the dark room the Hering-figure subtended 9° of arc to the eye, while in the distance-viewing it subtended angles of less than 50′. Comparison is thus confounded with differences arising from the inclusion of peripheral vision as against the use of central vision alone. Nevertheless, the very general theoretical account here offered in terms of different modes of perceptual patterning in the two situations does not need to be altered on this account. As is the case with night-vision, peripheral vision tends to proceed by aligning centers of density to form primary patterns under impoverished conditions.

Because of the limited range of values used, the present attempt to compare complexly different perceptual situations in terms of ability to resolve the same sizes of physical detail has resulted in considerable over-simplification. In the first place, the sizes of acuity-targets were chosen to require luminance barely extending beyond the range of scotopic vision, when viewed from a distance of 3 m.[9] Therefore, the gradual linear slope obtained in the dark room might well have changed to a steeper slope at the next stage of higher photopic luminances of the normal visual acuity curve, had smaller acuity-targets been included. Furthermore, keeping the same distance of viewing and using a wide range of luminances, a family of curves could be obtained by changing the size of the Hering-figure. Curves for the larger sizes might then be expected to have both scotopic and photopic sections of different slopes, but even the curves for smaller figures would be likely to have two slopes, not by reason of scotopic function, but because negative illusion appears to be derived from the segregation of proportionately larger sizes of detail than that required for normal illusion. The distance-viewing curve of Fig. 1 may thus be interpreted as a cross section of such a possible family of curves at the given level of photopic luminance.

The present findings suggest that Hering's illusion is primarily determined by local effects at intersections of lines and bars, but the data are insufficient to decide between the rival theoretical accounts. A general conclusion from these studies is that lesser magnitudes of the normal Hering-illusion can be achieved at certain stages of indefinite form, but that the normal amount of illusion of the typical demonstration requires resolution of detail of an order only possible with photopic vision.

---

[9] J. L. Brown, Effect of different preadapting luminances on the resolution of visual detail during dark adaptation, *J. opt. Soc. Am.*, 44, 1954, 48-55.

SUMMARY

The amount of illusion in a Hering-figure with radial lines and horizontal bars of equal width was determined for 12 *S*s at five levels of luminance in a dark room and from five distances of viewing along a corridor. Each level of luminance or distance used with each *S* was that at which *S* could just resolve one of five Snellen 'E' acuity-targets. Using the same procedure, 12 control *S*s set the bars parallel without the radial lines. Their settings did not differ significantly from parallel under any of the conditions. Although there was evidence that the radial lines were discriminable at all levels of impoverishment used, at some of these levels, in both scoptopic and photopic vision, the settings of the experimental *S*s did not differ from those of the controls. Interpretation of the comparative results with different modes but similar degrees of impoverishment was made difficult by a reversal of the normal illusion obtained with the smaller angular sizes of the figure in distance-viewing. Settings made in the dark room showed a gradual increase in the amount of illusion with improvement of conditions of viewing. The findings were interpreted as being broadly consistent with known facts of night- and day-vision, and were considered to favor the position that the amount of Hering-illusion is determined more by the influence of certain small regions of the figure than by any effects arising from its over-all patterning.

THE EFFECT OF DISCRIMINATIVE TRAINING ON THE GRADIENT OF STIMULUS-GENERALIZATION

By ERIC G. HEINEMANN, Brooklyn College,
and ROCHELLE L. RUDOLPH, Vassar College

The training procedures used in attempts to bring behavior under environmental control fall into two classes. In non-differential training, a selected response is simply reinforced repeatedly in the presence of the chosen stimulus. In differential or discriminative training, two stimuli are used. The response is reinforced in the presence of one stimulus, but not in the presence of another. The 'precision' of the stimulus-control established is indicated by the shape of the gradient of stimulus-generalization. A horizontal gradient covering the entire stimulus-range indicates that the stimulus has no effect on the behavior.

Lashley and Wade argued that non-horizontal gradients of stimulus-generalization arise only as a result of discriminative training.[1] This question has been the subject of a number of investigations, but there are some discrepancies among the results that remain to be explained.

Several experiments have yielded extremely clearcut support for Lashley's basic position. Ferster showed that the presence or absence of two stimuli (a steady sound and a slowly flickering light) that had been continuously present while behavior was reinforced had no effect on responding during extinction.[2] Similarly, Jenkins and Harrison, who studied generalization along the dimension of frequency of a tone, found gradients of generalization which were virtually horizontal lines when no differential training was given.[3] Sloping gradients appeared as a consequence of differental reinforcement based on the presence or absence of the tone; training to discriminate between two non-zero values of the stimulus was not necessary.

MacCaslin, Wodinsky, and Bitterman, using a Lashley jumping stand, pre-trained their *Ss* to avoid the negative stimulus with the result that, although the *Ss* had not previously encountered the positive stimulus, they made very few unreinforced responses to the negative one.[4] Under these conditions, the positive stimulus failed to gain any control over responding, in spite of regular reinforcement. All of these experiments support the conclusion stated by Ferster that a "crucial event in establishing discriminative control over behavior is the nonreinforced occurrence of the response in the absence of the positive stimulus."[5]

---

* Received for publication March 23, 1962.

[1] K. S. Lashley, and Marjorie Wade, The Pavlovian theory of generalization, *Psychol. Rev.*, 53, 1946, 72-87.

[2] C. B. Ferster, The effect on extinction responding of stimuli continuously present during conditioning, *J. exp. Psychol.*, 42, 1951, 443-449.

[3] H. M. Jenkins and R. H. Harrison, Effect of discrimination training on auditory generalization, *ibid.*, 59, 1960, 246-253.

[4] E. F. MacCaslin, Jerome Wodinsky, and M. E. Bitterman, Stimulus-generalization as a function of prior training, this JOURNAL, 65, 1952, 1-15.

[5] Ferster, *op. cit.*, 448.

There is, however, some evidence that generalization gradients of non-zero slope can be obtained without explicit differential training.[6] The experiment by Grant and Schiller is unique in that it deals with a classically conditioned response. Differential reinforcement based on the presence or absence of the *CS* is, of course, an intrinsic feature of the classical conditioning procedure. There is thus no conflict between the findings of Grant and Schiller and those cited earlier. The reasons for the remaining discrepancies are less clear. A detailed examination of the experiments by Grice, by Schlosberg and Solomon, and by Guttman and Kalish suggests that in each of these experiments some incidental discriminative training may have occurred. What is involved can be made clear by a comparison of the experiment of Guttman and Kalish with that of Jenkins and Harrison. The methods used in these experiments appear to have been sufficiently similar to justify the conclusion that the difference in the results of these two experiments must stem from the use of visual stimuli in one and auditory stimuli in the other. This is the conclusion reached by Jenkins and Harrison. They also suggested the possible importance of the fact that "the visual stimulus appeared directly on the response key, whereas the auditory stimulus was diffuse and probably unlocalized."[7]

It should be noted that the presence or absence of a small visual stimulus on *S*'s retinas depends on its orientation in the apparatus, while there is no comparable dependence on orientation for the auditory stimulus. With reference to pigeons in a Skinner box, it seems very likely that, if the visual stimulus is presented on the response-key, so-called nondifferential training methods actually do involve some differential training, at least in the sense of training to discriminate between the presence and absence of the stimulus. Such differential training occurs most obviously while the animal is being taught to peck the key by the method of successive approximation. During this phase of training, behavior such as holding the head at the appropriate height, moving the head forward, and so forth, is reinforced when it occurs near the key (when the visual stimulus is likely to be imaged on *S*'s retinas), while similar behavior goes unreinforced if it occurs in some place other than the immediate vicinity of the key (when the visual stimulus is much less likely to be imaged on the *S*'s retinas). It is obvious that such differential training could not occur if the visual stimulus completely surrounded the *S*, as in the experiment by Ferster. Furthermore, it would seem that the smaller the area of the stimulus, the smaller would be the likelihood of its being present on the *S*'s retinas during the occurrence of any given unreinforced sequence of behavior. Hence, the amount of differential training that occurs should be related to the geometric size of the stimulus.

It was the purpose of the present experiment to explore the possibility that, with visual stimuli, the steepness of the gradient of generalization is determined by the area of the stimulus used. Generalization-gradients along the dimension of luminance were measured for stimuli of three different areas by methods generally similar to those used by Guttman and Kalish.

---

[6] D. A. Grant and J. J. Schiller, Generalization of the conditioned galvanic skin responses to visual stimuli, *J. exp. Psychol.*, 46, 1953, 309-313; G. R. Grice, The acquisition of a visual discrimination habit following responses to a single stimulus, *ibid.*, 38, 1948, 633-642; Norman Guttman and H. I. Kalish, Discriminability and stimulus generalization, *ibid.*, 51, 1956, 79-88; Harold Schlosberg and R. I. Solomon, Latency of response in a choice discrimination, *ibid.*, 33, 1943, 22-39.

[7] Jenkins and Harrison, *op. cit.*, 252.

Subjects. The *Ss* were 18 experimentally naïve pigeons. They were maintained by restricted feeding at 80% of their normal weight.

Apparatus. A modified version of the Skinner automatic key-pecking apparatus was used. The *S*'s compartment was a box 16 in. high, 12 in. wide, and 19 in. long (interior dimensions), constructed of wood and hardware cloth. In other respects the general features of the apparatus were the same as those described by Ferster and Skinner.[8] The reinforcement was a 2.5-sec. period of access to a tray of mixed grain. The roof of the box was made of hardware cloth. A single frosted bulb suspended above the roof provided constant illumination throughout the experiment.

The visual stimuli were made of 'white,' 'light gray' or 'dark gray' cardboard. When mounted in the apparatus, these cardboard surfaces had luminances of 15.4, 10.1, and 1.5 ft.-L., respectively. Three sets of stimuli differing in area were used. Each set consisted of three stimuli, one for each of the luminances listed, but all of the same area.

The largest stimulus used consisted of a cardboard that covered the entire 12 in. × 16 in. response-panel of the apparatus, except for two holes. Approximately in the center of the cardboard, there was a circular hole 1 in. in diameter, behind which the response-key was located. Somewhat below this was a rectangular hole, 2.4 in. × 2.1 in., which permitted access to the food-tray. The response-key was made of nonselectively reflecting plastic having a luminance under the conditions of the experiment of approximately 9 ft.-L.

The middle-sized stimulus consisted of an annulus, centered about the response-key, that had an inner diameter of 1 in. and an outer diameter of 5¼ in. The small stimulus, also an annulus centered about the response-key, had an inner diameter of 1 in. and an outer diameter of 1¾ in. When the middle-sized or small stimuli were in use, the rest of the response-panel was covered by 'black' cardboard having a luminance of 0.41 ft.-L.

Procedure. The *Ss* were divided into three groups of six. For each group a stimulus of 10.1 ft.-L. was exposed continuously during all phases of training, but its area was different for each group. Group *L* was trained with the large stimulus; Group *M* with the middle-sized stimulus; and Group *S* with the small stimulus.

After preliminary training to eat from the magazine, the *Ss* were trained to peck the key by the method of successive approximation. Key-pecking was reinforced regularly for several 20-min. sessions, and the *Ss* were then placed on a variable-interval (*VI*) schedule with a mean interval of 1 min. Training on the *VI*-schedule was continued until a predetermined criterion of response-stabilization was attained; namely, two cumulative-response plots covering 20-min. periods, obtained on successive days, consisting of straight lines showing no measurable differences in slope.

In the tests for generalization, each *S* was exposed only to stimuli of the same area as the one used in its training. The luminances of the stimuli used in the generalization-tests were 1.5, 10.1, and 15.4 ft.-L. Only a single luminance was presented in each testing session, and there was one testing session for each luminance. The number of responses during the 10-min. period following the first peck was recorded. For each group, the order of presentation of the stimuli was arranged according to a matrix design in which each stimulus appeared in each ordinal position equally often. Responses were not reinforced during the tests for generalization. Between the successive pairs of test, a retraining session was given during which the original stimulus was present and responses were reinforced on the *VI*-schedule.

---

[8] C. B. Ferster and B. F. Skinner, *Schedules of Reinforcement,* 1957.

Results. Fig. 1 shows the percentage of responses made to each stimulus by each group of *S*s. Each point is a mean based on data for six *S*s. The gradient for the largest stimulus is virtually flat. The other two stimulus-conditions give rise to sloping gradients, the smaller stimulus yielding the steeper slope.

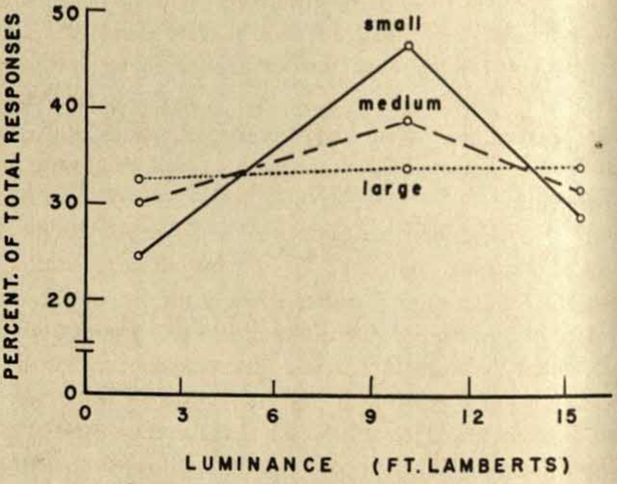


FIG. 1. GENERALIZATION-GRADIENTS OBTAINED AFTER TRAINING WITH LARGE, MEDIUM, AND SMALL STIMULI OF 10.1 FT.-L.

A three-way analysis of variance was done on the total number of responses made by each *S* in each test period. The variables were individuals, stimulus-area, and stimulus-luminance.[9] The effect of area was not significant, indicating that there were no substantial differences among the three groups in total number of responses. Highly significant effects were found for luminance ($F = 34.16$ with 2 and 30 *df*, $p < 0.001$) and for the interaction between area and luminance ($F = 10.41$ with 4 and 30

---

[9] Quinn McNemar, *Psychological Statistics*, 1955, 332.

df, $p < 0.001$). The significant interaction between area and luminance makes it possible to reject the hypothesis that the three curves shown in Fig. 1 have the same shape. Thus, the sum of the responses made during the three extinction-runs (generalization-tests) is about the same for the three groups, but the way in which these responses are distributed among the three stimulus-conditions differs significantly.

To discover for which of the groups there was a significant dependence of response-frequncy upon luminance during the generalization-tests, a separate analysis of variance (luminance *vs.* *S*s) was made for each group. A significant effect of luminance was found for Group *S* ($F = 24.4$ with 2 and 10 *df*, $p < 0.001$) and for Group *M* ($F = 9.81$ with 2 and 10 *df*, $p < 0.01$), but not for Group *L* ($F = 1.20$ with 2 and 10 *df*, $p > 0.05$).

Discussion. The results show clearly that the shape of the generalization-gradient for the dimension of luminance depends on the area of the stimuli used in training. With the very largest stimuli the gradient is a horizontal line, indicating that the stimulus failed to acquire any control over the behavior.

There can be little doubt that Groups *M* and *S* received some differential training in the manner described in the introductory section. It seems likely also that these groups differed systematically in the 'amount' of differential training received. By difference in the 'amount' of differential training is meant the difference in the relative frequencies with which the stimulus accompanied reinforced and unreinforced responses. Since the stimuli of different areas are all present near the response-key, the proportion of reinforced responses occurring in the presence of the stimulus should be independent of area, but the proportion of unreinforced responses occurring in the (retinal) presence of the stimulus should decrease directly with geometric size. Such differences in the differential training received by the three groups may well be sufficient to account for the differences in the steepness of the gradients of generalization that were found.

If this analysis is generally correct, it would appear that some discriminative training must inevitably occur in any situation in which the stimulus is a fairly small visual area and the required response is a movement that is directed with respect to this stimulus, such as pecking at the stimulus, or walking or jumping through a door upon which the stimulus appears, or turning at a choice-point marked by the stimulus. It seems likely that the results of previous experiments which appeared to show that a stimulus does acquire some control over a response in the absence of differential training reflect some unintended differential training of this sort.

SUMMARY

Generalization along the dimension of luminance of a visual stimulus was measured by means of an operant procedure. The principal purpose was to compare the effects of training with stimuli of different geometric size. Three groups of pigeons were trained to peck at a response-key in the presence of a stimulus of 10.1 ft.-L. For one group, the stimulus was small; for the second group, it was intermediate in size; and for the third group it was a large rectangular surround. For each group, the stimuli used to test for generalization were of the same size and shape as that used in training. The gradient of generalization following training with the large stimulus was approximately a horizontal line. The other two stimulus-conditions gave rise to sloping gradients, the smaller stimulus yielding the steeper slope. It is suggested that differences in the slopes of the gradients arise from differences in the amount of discriminative training received by the three groups. This interpretation is based on the assumption that during the course of nominally 'nondifferential' training *S* is trained to discriminate the stimulus from its surroundings, and that the amount of discriminate training received is related to the geometric size of the stimulus.

ACCOMMODATION AS A DISTANCE-CUE

By HANS WALLACH, University of Oxford, and
CYNTHIA M. NORRIS, Swarthmore College

There are a number of perceptual properties which are multiply determined—that is, which are either caused or affected by more than one stimulus-condition. Two conditions of stimulation, for instance, determine perceived verticality; there are two that cause perceived visual motion, three that produce visual depth, and at least three that have an effect on perceived distance. When we wish to establish the effectiveness of one of such interacting conditions of stimulation, we attempt to construct a situation from which the others are eliminated or in other ways controlled. It seems desirable, however, to go beyond the mere proof of effectiveness and to discover the basis of the interaction.

A first step in the pursuit of this problem may be made by creating artificial viewing conditions under which one of a pair of cues is altered and thus comes into conflict with its normal partner. Under natural conditions each one of the paired cues tends to produce the particular perceptual property at the same value. If one of the cues is artificially changed to give a different value of the property, the question arises of how this conflict is resolved. We have found that under such conditions, one of the paired cues may dominate perceptual experience permanently, or the two cues may do so alternately, or the value of the property may represent a compromise between the conflicting cues. The work to be here reported grew out of such investigations. Dominance of one or the other of the conflicting cues, or in the case of compromise, the actually obtained value of the property, permits inferences to be drawn as to the relative strength of the paired cues.

Our experiment was concerned with two conditions of stimulation that determine perceived distance of an object in monocular vision, namely, perspective and accommodation. Of the two, accommodation usually has been considered as rather unimportant. By causing accommodation to contradict cues of perspective, the capacity of accommodation to compete with cues of perspective can be explored.

The following observations led to the plan of the experiment to be reported.

---

* Received for prior publication July 28, 1963. This work was supported by a grant from the National Science Foundation.

They were made with an arrangement which presents strongly deceptive cues of perspective. This arrangement consisted mainly in an oblique projection of a cube with 7.5 cm. length of edge on a horizontal plane (Fig. 1). The three faces shown in the drawing were painted three different tones of gray. To assure the sharpest possible face-edges, three panels in the shapes and colors of the three faces were cut out of cardboard and fitted together to form the cube-projection. This cube projection was placed on a table covered with a checkered oilcloth. A vertical cardboard screen was placed 31 cm. from the lower end of the cube-projection. A peephole of 1.4 cm. in diameter was cut in the cardboard at a height of 16.5 cm. above the table-surface. The location of the peephole was so chosen that an eye peering through the hole occupied the station-point of the projection. Since, viewed from this point, the cube-projection produced the retinal image of a cube as seen obliquely from above, it was at least possible that *O* would see a regular gray cube standing on the table in the location of the lower end of the cube-projection. Actually, this was only partly true. Some *O*s did report seeing a regularly shaped cube, but the majority described a three-dimensional form leaning somewhat backward; the two faces that in a regular cube would have been vertical appeared to be on a backward slant. Inasmuch as these two faces were objectively located in the horizontal table-surface, their observed slant represents a veering toward veridicality. It seemed to us that, under the prevailing conditions, accommodation was the factor most likely to be responsible for mediating an influence of the objective arrangement of the faces. This suspicion was confirmed when we equipped our *O*s with an artificial pupil. The slant toward the rear was no longer noticed, and all *O*s reported seeing a regularly shaped cube.

This effect of accommodation is surprising on two counts. In altering the shape of the perceived three-dimensional object caused by the cube-projection, accommodation shows an unexpected effectiveness, and it operates as a cue not merely to distance but also to form. It is most likely that the reason why a pattern of stimulation in the shape of a cube-projection leads to the perception of a cube lies in the very high incidence of experience with orthogonal visual objects. Thus, the very cause of the perception of the cube projection as a solid should tend to make it a cube and nothing but a cube. Any factor in the situation which causes the solid to have slanting faces must, therefore, be of considerable potency. That accommodation operates here as a factor in form perception is obvious, inasmuch as a majority of *O*s upon removal of the artificial pupil report a change in the *shape* of the perceived solid and not merely a change in the distance of its top.

Our experiment is an attempt to establish the potency of accommodation as a distance-cue through measurement. The top face of the apparent solid seems to be closer when the cube-projection is viewed through an artificial pupil than when it is not. Measurement of the apparent distance of the top face under the two conditions would, therefore, reflect the effect of the artificial pupil and the influence of accommodation, inasmuch as the pupil that was used was fully capable of eliminating accommodation as a cue for distance. Rather than measuring distance as such, we used perceived size as an indicator. It is well known that, with adequate cues for distance, the perceived size of an object depends on the size of its retinal image and its distance from O. With image size kept constant, perceived size is proportional to this distance. Size-perception thus can be used to determine how accurately the given distance-cues represent that distance. An object so

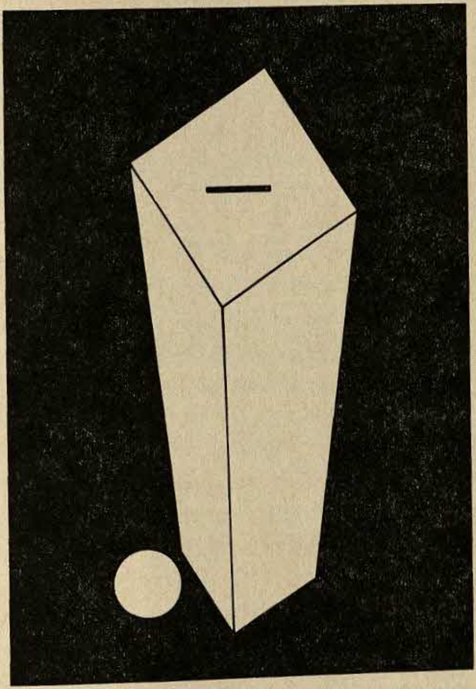


FIG. 1. OUTLINE DRAWING OF THE CUBE-PROJECTION ON A DARK BACKGROUND

placed on the cube-projection that it appears to lie on the top face of the perceived solid (see heavy line in Fig. 1) thus should be seen somewhat larger when the solid appears to slant backward. The apparent sizes of the object under the two conditions can be measured by having each matched by a similar object of variable size. In addition to making these two measurements, we also obtained a match for the apparent size of the object without the deceptive distance-cues produced by the cube-projection. A checkered oilcloth exactly like the one on which the cube-projection was resting was spread over it and the rest of the table surface, and the object was put back in its old position. With a still greater apparent distance, the object now should appear still larger than before.

PROCEDURE

Our objects were black cardboard squares of various sizes which were kept in an upright position by little wooden blocks glued to their backs flush with one edge. The standard, 5 cm. on a side, was put in the same place under all three conditions of the experiment, namely, in the center of that part of the cube-projection which forms the top face of the perceived cube. The position of the comparison-object was so chosen that it actually was at the same distance from the eye and approximately at the same elevation above the table surface at which the standard square would have been if the cube-projection were replaced by the cube from which it was derived. The place for the comparison-squares was on top of a brass cylinder 6.8 cm. high which stood on the table close to the near end of the cube-projection (white disk in Fig. 1). The standard square was 66 cm. and the comparison square 38 cm. from *O*'s eye. The size of the comparison-squares varied in steps of 2 mm.

The artificial pupil consisted of a disk of sheet metal with a hole of 0.95 mm. diameter in its center. The aperture was large enough to allow simultaneous viewing of the standard and comparison-squares. According to Campbell, such a pupil yields a depth focus of 1.65 diopters, large enough to bridge the difference between 66 and 38 cm. viewing distance which involves a change by 1.1 diopters.[1]

Matches were obtained by the method of limits with three lower and three upper limits taken. The average of each *O*'s six first equality-judgments became his score. When any two of these six judgments differed by 0.8 cm. or more for any one of an *O*'s matches, his performance was held inadequate and his scores were not used. Three *O*s were thus eliminated and the means to be reported were for 12 *O*s. First, *O*s made matches for the two conditions with the cube-projection visible; the condition in which the cube-projection was covered always came last. For the two cube-conditions the sequence was varied; alternate *O*s started by using the artificial pupil.

RESULTS

As was expected on account of the difference in appearance of the cube-projection with and without the artificial pupil, the size-matches obtained *with* the pupil were such as to imply a shorter distance of the standard square from the eye than those obtained *without* the pupil. When accommodation was effective, it counteracted the effect of the cube-projection which tends to bring the standard square closer to the eye. The standard square appeared smaller with the pupil than without it. The mean matches obtained were 3.26 cm. *with* the pupil and 3.68 cm. *without* it, the difference of 0.42 cm. being reliable at the 0.001 level ($t = 5.87$). With the objective size and distance of the standard square taken as the base, these size-matches imply apparent distances of 43.0 cm. and 48.5 cm. between the standard square and *O*. The difference of 5.6 cm. is the effect of accommodation on distance that is implied by the difference in apparent size.

---

[1] F. W. Campbell, The depth of field of the human eye, *Optica Acta,* 4, 1957, 157.

The effect of the cube-projection as such is demonstrated by the comparison of these mean sizes with a mean size match of 4.17 cm. obtained when the cube-projection was covered with a tablecloth. It is interesting that the difference of 0.49 cm. between this value and the mean obtained with the cube-projection and without the pupil is not much larger than the further size reduction of 0.42 cm. produced by employing the pupil. It also is less reliable ($t = 3.38$). This is due to the fact that the variability of the individual matches was much higher for the tablecloth-condition. The individual scores here showed an *SD* of 0.67 cm. whereas the *SDs* for the cube-conditions were 0.20 cm. with pupil and 0.25 without. This difference is not explained by the greater apparent distance of the standard square in the tablecloth-condition; its mean apparent size implies a distance of only 55.0 cm. Inasmuch as the difference between the *SDs* of the tablecloth-condition on the one hand and the cube condition without pupil on the other is highly reliable ($F = 7.0$), we can say that the cube-conditions have a more homogeneous effect on the *O*s than the tablecloth-condition, and seem to provide the more definite cue for distance.

DISCUSSION

Our size-matches showed that accommodation can function as a potent cue for distance. By reducing the apparent distance of the standard square, the cube-projection caused the apparent size of the standard to be reduced from 4.17 cm. (tablecloth condition) to 3.26 cm. (with artificial pupil), a change of 28%. When accommodation was introduced by omission of the pupil—providing a potential cue for the veridical distance of the standard—its apparent size was reduced to 3.68 cm., a change amounting only to 15%. The effect of the cube projection was reduced 45% by having accommodation counteract it.

The latter finding will come as a surprise to many. In summarizing the early work on the effect of oculomotor adjustment on the perception of distance, Woodworth states that "all the evidence favors the great predominance of visual cues" over oculomotor adjustment.[2] A recent experiment by Heinemann, Tulving and Nachmias on the effect of accommodation on apparent size also seems to lead to a different view of the effectiveness of accommodation.[3] These authors obtained size-matches with accommodation the sole cue for distance and found not more than 7% size-change for a distance-change equivalent to one diopter of accommodation. In our experiment, a distance-change equivalent to 1.1 diopters caused a size-change of 13%, and this effect is achieved in the process of counteracting perspective cues to distance. The discrepancy is, however, readily explained. The earlier authors created conditions unfavorable to demonstrating an effect of accommodation as a cue for distance by placing the variable object at a distance of 4 m. from *O*. At such a large distance, accommodation cannot be expected to operate effectively. If accommodation can be assumed to operate with an error of ±0.25 diopters—Campbell found this value for the detection of blurring for a pupil diameter of 4.7 mm.[4]—the range over which distance remains undetermined when the given distance is 4 m. reaches from infinity on one side to a distance of 2 m. on the other. Our results suggest that a stronger effect on size would have been obtained if both comparison-objects had been located within the range of high effectiveness of accommodation, that is, of distances up to 1 m.

---

[2] R. S. Woodworth, *Experimental Psychology*, 1938, 680.

[3] E. G. Heinemann, Endel Tulving, and Jacob Nachmias, this JOURNAL, 72, 1959, 32-45.

---

[4] Campbell, *op. cit.,* 160.

CONDITIONING AND EXTINCTION IN THE PLANARIAN

By RONALD BAXTER and H. D. KIMMEL, University of Florida

McConnell and his colleagues have reported successful conditioning and retention of conditioning after regeneration in the two cut halves of the Planarian, *Dugesia dorotocephala.*[1] The novelty and potential theoretical significance of this work make it imperative that the conclusion be securely established that conditioning actually occurred rather than some non-associative phenomenon. There is always the possibility that what seems to be conditioning may, upon closer inspection, turn out to be pseudo-conditioning or an effect of sensitization.

The primary purpose of this experiment, therefore, was to repeat the procedure employed by these investigators and, at the same time, to run an unpaired control for comparison with the experimental condition. In addition, the intensity of both the conditioned and unconditioned stimuli were varied to determine whether the effects of these variables would conform to those obtained in earlier studies on classical conditioning.[2] Since no data on extinction of conditioned responses (*CR*s) in the planarian have yet been reported, we decided, furthermore, to follow conditioning with standard extinction-trials, rather than to cut the animals in half and observe retention after regeneration.

Method: (1) Subjects. Twenty-four planaria (*Dugesia tigrina*), varying in length from 8 to 17 mm., were obtained from a local pond. The *S*s were kept in individual circular dishes. The water in these dishes and in the conditioning trough was obtained from the same pond as the animals.

(2) Apparatus. The conditioning trough was constructed from a sheet of lucite, $4\frac{3}{4}$ in. long, depressed longitudinally into a trough $2\frac{1}{3}$ in. wide and $1\frac{1}{4}$ in. deep. Copper plates forming the two ends of the trough also served as the electrodes which delivered the unconditioned stimulus (*UCS*). The trough was filled with pond water.

Two *UCS* intensities were produced by placing either 30,000 or 40,000 ohms resistance in series with the copper electrodes and a 6-v. storage battery, the lower value being used to produce the "strong" *UCS* and the higher one to produce the 'weak' *UCS*. Two conditioned stimulus-intensities (*CS*) were produced by

---

* Received for publication April 9, 1962.

[1] J. V. McConnell, A. L. Jacobson, and D. P. Kimble. The effects of regeneration upon retention of a conditioned response in the planarian, *J. comp. physiol. Psychol.*, 52, 1959, 1-5; Robert Thompson, and J. V. McConnell, Classical conditioning in the planarium, *Dugesia dorotocephala, ibid.*, 48, 1955, 65-68.

[2] Gregory A. Kimble, *Hilgard and Marquis' Conditioning and Learning*, 1961, 44-77.

using either a 25-w. or 100-w. bulb, located 5 in. above the center of the trough. Onset and offset of the stimuli were controlled manually by *E*.[3]

(3) Design. The *S*s were assigned randomly to the eight cells of a 2 × 2 × 2 factorial design, with three *S*s in each cell. The eight cells represented all possible combinations of levels of the three independent variables. The independent variables were: (a) *CS-UCS* contiguity (*i.e.* paired or unpaired); (b) *CS* intensity (weak or strong); and (c) *UCS* intensity (weak or strong).

(4) Procedure. The *S*s were run individually in a quiet, dimly lighted room. Prior to beginning each daily block of trials, *S* was transferred to the conditioning trough in a glass pipette and given 5 min. to adapt to the trough. Each *S* received 50 paired or unpaired *CS-UCS* trials per day for 5 days and 50 extinction trials on the sixth day. Individual trials were begun only when the animal was moving in the longitudinal plane of the trough toward one of the copper electrodes, since the *UCS* was found to be ineffective otherwise.

The duration of the *CS* was 3 sec., controlled by *E*'s counting, "one Mississippi, etc." The duration of the *UCS* was 1 sec., similarly determined. On paired conditioning trials, the *UCS* shared the last 1 sec. of the *CS*.

The intertrial interval was always at least 30 sec. in the groups which received paired *CS-UCS* presentations. In the unpaired groups either a *CS* or an *UCS* was delivered first, the choice being random, and the other stimulus was presented 15 sec. later. Then at least 15 sec. elapsed before the next trial was begun. Extinction-trials were separated by at least 30 sec. for all *S*s.

Results. A *CR* was defined as a longitudinal contraction of the animal's body which occurred within 2 sec. of the onset of the *CS, i.e.* before the onset of the *UCS*. Fig. 1 presents the mean number of *CR*s per block of 50 conditioning trials for the combined paired groups and the combined unpaired groups. The figure reveals a steady divergence of the two curves.

Analysis of variance of the total number of *CR*s during conditioning showed that the main effects of all three independent variables were significant beyond the 1% level (error variance = 243.3) but none of the interactions approached significance. The differences in mean number of *CR*s during conditioning due to both *CS* and *UCS* intensity favored the stronger stimulus.

Table I presents the means and standard deviations of the total number of *CR*s made during extinction. As is indicated in the table, the difference which had been found during conditioning between the paired and unpaired conditions was absent during extinction. The differences due to the intensity of the *CS* and *UCS* continued to be present in the same direction as during conditioning.

Analysis of variance of the total number of extinction-*CR*s showed that the effects of the intensity of the *CS* and *UCS* were significant beyond the 1% level and the interaction between *CS-UCS* contiguity and the intensity of the *CS* was significant at the 5% level (error variance = 7.7). The significant interaction was due to the fact that the effect of the intensity of the *CS* during extinction was confined primarily to the unpaired groups.

---

[3] J. V. McConnell, P. R. Cornwell, and Margaret Clay, An apparatus for conditioning planaria, this JOURNAL, 73, 1960, 618-622.

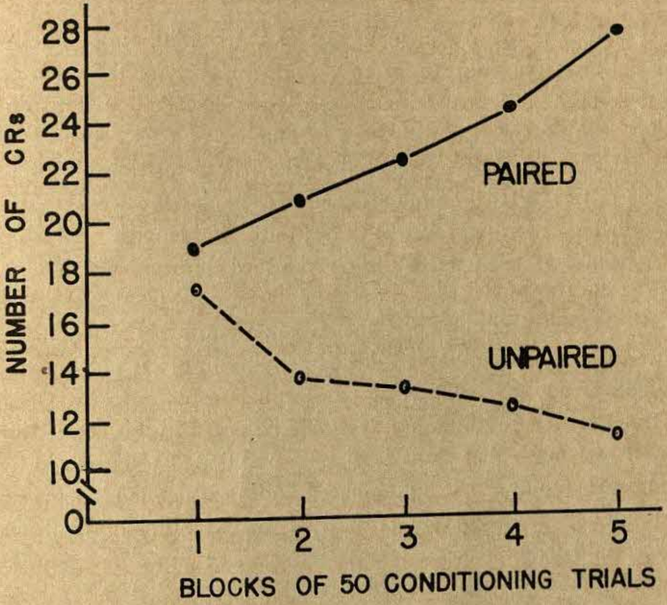


FIG. 1. MEAN NUMBER OF *CRs* IN SUCCESSIVE DAILY BLOCKS OF 50 CONDITIONING TRIALS FOR THE PAIRED AND UNPAIRED GROUPS ($N = 12$ in each group)

TABLE I

MEANS AND STANDARD DEVIATIONS OF TOTAL NUMBER OF EXTINCTION-*CRs*

	Strong *UCS*			Weak *UCS*			Total		
Contiguity	Strong *CS*	Weak *CS*	Total	Strong *CS*	Weak *CS*	Total	Strong *CS*	Weak *CS*	Total
Paired:									
$\bar{X}$	11.3	12.7	12.0	11.0	7.0	9.0	11.2	9.8	10.5
SD	2.5	1.9		1.6	2.2				
Unpaired:									
$\bar{X}$	18.7	10.3	14.5	12.0	6.0	9.0	15.3	8.2	11.8
SD	1.3	2.6		3.6	2.4				
Total $\bar{X}$	15.0	11.5	13.2	11.5	6.5	9.0	13.2	9.0	11.1

Because of the fact that the significant difference between the paired and unpaired groups which had been observed during conditioning was absent in the data obtained during extinction, it was of interest to examine the number of extinction-*CRs* in blocks of 10 trials for the paired and unpaired groups. This comparison showed that the paired group dropped in number of *CR*s to the level of the unpaired group within the first 10 extinction-trials and that the two groups declined at roughly the same rate throughout the remainder of the extinction-trials.

Discussion. The consistent divergence of the conditioning curves of the paired and unpaired groups may be taken as support for the assertion that classical conditioning occurs in the planarian. There appears to be little basis for arguing that the difference in number of *CR*s between the two groups was due to anything other than the paired relationship of the *CS* and *UCS* in the paired group and its absence in the unpaired group.

Nevertheless, attention must be drawn to the suddenness with which the high frequency of response to the *CS* vanished during extinction. Almost as soon as the *UCS* was removed the paired groups dropped to the response-level of the unpaired groups. The question must be raised, therefore, whether the difference observed during conditioning was not due to *alpha* conditioning,[4] *i.e.* sensitization of the original response to the *CS,* rather than 'true' classical conditioning. It should be noted in this connection that Cornwell has shown that planaria give unconditioned responses to light in massed presentations of the light alone, a finding which lends some support to this notion.[5] Also relevant is the fact that McConnell, Jacobson, and Kimble found their evidence for the retention of conditioning in the regenerated sections of planaria which had been cut in half after conditioning by means of relearning savings scores rather than on tests using only the *CS*.[6] Since there is presently no immediately apparent way to differentiate between *alpha* and 'true' conditioned responses in this organism, the question remains unanswered.

Two factors must be considered in interpreting the significant effects of the intensity of the *CS* and *UCS* during conditioning and extinction. First, during conditioning neither of these effects interacted with the *CS-UCS* contiguity-variable. This is interpreted to mean that the intensity of the *CS* and *UCS* influenced performance rather than influencing habit-formation since the effect of each was equally present in the paired and unpaired groups. Secondly, both intensity-variables continued to exert significant effects during extinction, even though the *CS-UCS* contiguity-variable no longer exerted its influence. Again, the interpretation is made that the intensity of the *CS* and *UCS* must have influenced performance rather than habit formation.

---

[4] C. L. Hull, Learning: II. The factor of the conditioning reflex, in Carl Murchison (ed.), *A Handbook of General Experimental Psychology,* 1934, 431.

[5] P. R. Cornwell, A preliminary study of pseudo-conditioning under conditions of massed stimulation in the planarian, *The Worm Runner's Digest,* 2, 1960, 101-108.

[6] McConnell, Jacobson, and Kimble, *op. cit.,* 5.

SUMMARY

This study compared paired and unpaired *CS-UCS* in the planarian at two intensive levels each of the *CS* and UCS. Twenty-four animals were divided into the eight subgroups of the 2 × 2 × 2 factorial design. The *CS* was either a weak or strong light of 3-sec. duration and the *UCS* was either a weak or strong electric shock of 1-sec. duration. The *S*s in the paired conditions received the *UCS* during the last 1 sec. of the *CS*, while those in the unpaired conditions received the stimuli at least 15 sec. apart. Each *S* had 50 trials of paired or unpaired *CS-UCS* per day for 5 days and 50 extinction-trials on the sixth day. The dependent variable was the number of *CR*s, a *CR* being a longitudinal contraction of the animal's body which occurred within 2 sec. after the onset of the *CS*.

The curves showing the frequency of the responses of the paired and unpaired groups diverged increasingly throughout conditioning. This difference was statistically significant as were those which favored the strong *CS* over the weak *CS* and the strong *UCS* over the weak *UCS*. None of the interactions was significant. The difference between the paired and unpaired groups which was observed during conditioning vanished, however, during the 50 extinction-trials. The difference due to the intensity of the *CS* and *UCS* were still present during extinction, although the former difference was largely confined to the unpaired group.

It was concluded that the difference in the frequency of the responses during conditioning between the paired and unpaired groups was due to the contiguous *CS-UCS* relationship which obtained during conditioning for the paired group but not for the unpaired group. Whether this difference could be taken as evidence for 'true' conditioning rather than *alpha* conditioning in the planarian could not be resolved, since the effect vanished as soon as the *UCS* was removed during extinction and it has been shown that light alone is capable of producing longitudinal contractions in this organism. The effects of the intensity of the *CS* and *UCS* did not interact with the *CS-UCS* contiguity-variable and were, therefore, assumed to have influenced performance rather than conditioning.

THE ROLE OF THE INTERPOLATED STIMULUS IN PRODUCING KINESTHETIC FIGURAL AFTER-EFFECTS

By C. G. COSTELLO, BEATRIZ HERRERA, and H. C. HOLLAND,
University of London

Earlier workers on kinesthetic figural after-effects (*KFAE*) discussed them in terms of satiational theory.[1] The degree of the after-effect was considered to be proportional to the degree of satiation which in turn, was proportional to the duration of stimulation by the inspection-block.

Broadbent and Northrup, however, independently found that pseudo-after-effects occurred after a number of trials when an inspection-block was not interpolated.[2] They suggested, therefore, that the pseudo-after-effects resulted from shifts in the scale of judgment caused by repeated experience with the tapered comparison-block. Broadbent also suggested that the interpolated experience of the inspection-block may cause the forgetting of a previous series of judgments thus resulting in a greater bias in the judgments following the period of stimulation.

Costello is of the opinion that Broadbent relegates the inspection-block to a too minor role. He believes that the experience with the inspection-block has a direct effect on the frame of reference relating the test-block to the comparison-block rather than merely interfering with the frame of reference established by earlier trials.[3] This proposal is supported by the finding of Charles and Duncan that the amount of the after-effect first increased and then decreased as the difference in width between the test- and inspection-block increased.[4]

That after-effects are produced by the interpolated inspection-block through a change in the frame of reference relating the test-block to the comparison-block rather than by the production of cortical satiation has been suggested also by Northrup.[5] She found that after-effects produced by just holding the inspection-block were significantly greater than those produced by the standard satiation procedure of rubbing the inspection-block.[6] This finding would appear to be more in

---

* Received for publication May 4, 1962. From the Institute of Psychiatry, Maudsley Hospital.

[1] G. S. Klein and David Krech, Cortical conductivity in the brain injured, *J. Person*, 21, 1952, 118-148; Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesis, *Miscellanea Psychologia Albert Michotte*, 1947, 196-220; H. J. Eysenck, Cortical inhibition, figural after-effect and theory of personality, *J. abnorm. soc. Psychol.*, 51, 1955, 94-106.

[2] D. E. Broadbent, Psychophysical methods and individual differences in the kinesthetic figural after-effect, *Brit. J. Psychol.*, 52, 97-104; D. R. Northrup, The kinesthetic figural after-effect: Methodological problems, Unpublished Master's thesis, University of Illinois, 1960, 14-18.

[3] C. G. Costello, The effects of meprobamate on kinesthetic figural after-effects, *Brit. J. Psychol.*, 53, 1962, 17-26.

[4] J. P. Charles and C. P. Duncan, The distance gradient in kinesthetic figural after-effect, *J. exp. Psychol.*, 57, 1959, 164-170.

[5] Northrup, *op. cit.*, 33-34.

[6] Northrup, *op. cit.*, 23-26.

keeping with the frame of reference theory of after-effects than with the satiational theory. According to the latter theory, the degree of after-effect is proportional to the duration of stimulation by the inspection-block.

The frame of reference theory is supported also by the findings of Jaffe.[7] His subjects looked at parallel lines equal in width to the test block during estimation of the test-block width and wider or narrower than the test-block during stimulation on an inspection-block equal in width to the test-block. They showed significant distortions in judgment of the test-block width after the wider visual stimuli. Jaffe himself suggests that "the contrast between the visual objects resulted in a change in the general frame of reference and this in turn influenced the kinesthetic judgment."[8]

The frame of reference theory receives some support also in the findings of Natadze.[9] In one of his experiments, he had *S* imagine, with his eyes closed, that he was grasping two wooden balls with his hands, a large one in one hand and a small one in the other. After this was repeated for 15 times, two balls objectively equal in size were placed for comparison in the *S*'s extended hands. The majority of his *Ss* developed a contrast-illusion of perception—the ball on that side on which the smaller ball was visualized was perceived as being the larger one, and vice versa.

The theory would also seem to fit in better than the satiational theory with the permanency of the after-effects reported by Wertheimer and Leventhal since, according to the satiational theory, the after-effects are temporal phenomena that can be measured only before the cortical satiation dissipates.[10]

The 'frame of reference theory' would lead one to predict that the width of the inspection-block would be a more important variable than the duration of stimulation in determining the degree of the after-effect. Pilot work carried out in our laboratory, however, and comparison with other studies, particularly that of Krauskopf and Engen,[11] suggested the hypothesis that the amount of after-effect is a function of both the width of the inspection-block and the duration of stimulation. The study reported here was designed to test this hypothesis.

METHOD

Subjects. Thirty-six *Ss* (28 men and 8 women) were used; their ages ranged from 16 to 40 yr. They were divided randomly into nine groups of four *Ss* each.

Procedure. Each *S* was tested individually during one session. The duration of stimulation by the inspection-block and the size of the inspection-block for each of the nine groups were determined by a 3 × 3 factorial design. The sizes of the inspection-blocks were 2½, 3½ and 4 in. Durations of stimulation were 5, 30, and 60 sec.

---

[7] Robert Jaffe, The influence of visual stimulation on kinesthetic figural after-effects, this JOURNAL, 69, 1956, 70-75.

[8] Jaffe, *op. cit.*, 75.

[9] Rewaz Natadze, Emergence of set on the basis of imaginal situations, *Brit. J. Psychol.*, 51, 1960, 237-245.

[10] Michael Wertheimer and C. M. Leventhal, Permanent satiation phenomena with kinesthetic figural after-effects, *J. exp. Psychol.* 55, 1958, 255-257.

[11] John Krauskopf and Trygg Engen, The distance paradox in kinesthetic figural after-effects, this JOURNAL, 73, 1960, 298-301.

The width of the test-block was 1½ in. The comparison-block was ½-in. wide at the near end, increasing in width to 4 in. at the far end. The apparatus is illustrated elsewhere.[12]

S, blindfolded throughout the experiment, was seated at a table on which the test-block and the comparison-block were placed. The test-block and the inspection-block were always to the right of *S* and the comparison-block always to the left. The comparison-block remained in the same position with the narrower end toward *S*.

S was instructed to grip the comparison- and test-blocks, using the thumb and large finger of each hand. He was instructed to keep the hand holding the test-block still and to move the hand holding the comparison-block to find the width on the comparison-block that matched the width of the test-block.

Four pre-inspection trials alternately ascending and descending were given to find the point of subjective equality (*PSE*) and thus the constant error for each *S*.

The test-block was then replaced by one of the inspection-blocks and a period of stimulation was given at one of the levels listed above. *S* moved the thumb and large finger of his right hand back and forth along the sides of the inspection-block during the period of stimulation. Immediately afterward, *S* made four judgments of the width of the test-block, again using alternately ascending and descending trials.

RESULTS

The results for the descending and ascending trials were analyzed separately and in combination. The mean *KFAE*s (prestimulation estimates of the width of the test-block minus post-stimulation estimates) for the 36 *S*s were calculated as a function of duration of stimulation, of size of the inspection-block, and of the interaction between size and duration of stimulation. The means are presented in Table I.

It will be seen from an inspection of this table that there is a tendency for increases in the duration of stimulation to produce increases in *KFAE*s. The same tendency is not observed in relation to the sizes of the inspection-blocks.

Analyses of variance for both ascending and descending trials, and for the combined scores were carried out. The only significant source of variance in the three analyses is for "duration of stimulation" in the descending trials ($F = 3.39$ with 2 and 27 *df*).

Discussion. The absence of a significant interaction between duration of stimulation and size of the inspection-block is not in line with the hypothesis that the size of the *KFAE* is a function of both the width of the inspection-block and the duration of stimulation.

---

[12] Eysenck, *op. cit.*, 1955, 99.

The absence of a significant size-effect is not in line with the original hypothesis that the size of the *KFAE* is a function of the width of the inspection-block. On the other hand, the significant effect due to duration of stimulation for the descending trials appears to be more in line with the satiational theory than the frame of reference theory.

What can be said then about the position of the two theories at the present time? Data more in line with the frame of reference theory than with the satiational theory are: (1) Jaffe's finding of *KFAE*s produced by visual stimuli;[13] (2) Charles and Duncan's finding that the amount of the after-effect was a function of the difference in width between the test- and inspection-blocks;[14] (3) Northrup's finding that the *KFAE*s produced by just holding the inspection-block were significantly greater than those produced by rubbing the inspection-block,[15] (4) Natadze's report that phenomena similar to *KFAE*s are produced by imaginal situations,[16] and (5) the permanency of *KFAE*s reported by Wertheimer and Leventhal.[17]

TABLE I

MEAN *KFAE*S AS A FUNCTION OF DURATION OF STIMULATION, SIZE OF INSPECTION-BLOCK, AND THE INTERACTION BETWEEN DURATION OF STIMULATION AND SIZE

Duration of stimulation	Size of inspection-block											
	2½ in.			3½ in.			4 in.					
	A	D	C	A	D	C	A	D	C	$\bar{X}$A	$\bar{X}$D	$\bar{X}$C
5 sec.	0.55	1.75	1.10	0.88	1.50	1.19	1.50	0.35	0.93	0.98	1.20	1.08
30 sec.	1.60	1.48	1.04	0.60	1.30	0.95	1.25	2.03	1.64	1.15	1.60	1.21
60 sec.	1.70	2.93	2.32	1.28	1.77	1.53	1.45	4.10	2.78	1.48	2.93	2.21
	$\bar{X}$A	$\bar{X}$D	$\bar{X}$C	$\bar{X}$A	$\bar{X}$D	$\bar{X}$C	$\bar{X}$A	$\bar{X}$D	$\bar{X}$C			
	1.26	2.08	1.49	0.92	1.53	1.23	1.40	2.16	1.79			

A = Ascending trials; D = Descending trials; C = Combined trials.

Our finding that the size of the *KFAE*s is not a function of the size of the inspection-block or of the interaction between size and duration of stimulation, though not diametrically opposed to the frame of reference theory, does not lend it support.

Data favoring the satiational theory rather than the frame of reference theory are: (1) The finding reported in this paper that *KFAE*s increase with an increase in time stimulation, and (2) the finding that some recovery from the effects of the interpolated stimulation period does occur if a rest period is given.[18]

---

[13] Jaffe, *op. cit.*, 1956, 73.

[14] Charles and Duncan, *op. cit.*, 1959, 166-167.

[15] Northrup, *op. cit.*, 1960, 21-26.

[16] Natadze, *op. cit.*, 1960, 237-238.

[17] Wertheimer and Leventhal, *op. cit.*, 1958, 257.

More generally, it can be said that *KFAE*s can be produced in situations accounted for more readily by the frame of reference theory (visual stimulation, imaginal situations, short periods of stimulation, holding the inspection-block) or in situations accounted for more readily by satiation theory (rubbing an inspection-block). But differences in the size of *KFAE*s have been produced only by variations in stimulation accounted for more readily by satiation theory (differences in time of stimulation).

In conclusion, it would seem likely that both theories may be needed to account for kinethetic figural after-effects. In other words, the after-effects are a result of both cortical satiation and a shift in the frame of reference relating the test-block to the comparison-block. It is quite clear, however, that a considerable amount of experimental work must be done before an adequate account of the after-effects can be given.

SUMMARY

Kinesthetic figural after-effects (the apparent reduction in width of a block of wood after a period of stimulation by a wider block of wood) are discussed in terms of two theories—satiational theory and a frame of reference theory. The hypothesis that the size of *KFAE*s is a function of width of the inspection-block was formulated on the basis of the frame of reference theory, but was not supported by pilot experiments. The experiment reported was designed to test the hypothesis that the size of *KFAE*s is a function of both the width of the inspection-block and the duration of stimulation. The hypothesis was not supported by the experimental findings.

It was concluded, on the basis of the present findings and the findings of other investigators, that both the satiational theory and the frame of reference theory are needed to account for *KFAE*s.

---

[18] Eysenck, *op. cit.*, 1955, 102; Costello, *op. cit.*, 1962, 21.

RECOGNITION-TESTS, REPETITION, AND ALL-OR-NONE LEARNING

By JAMES G. MARTIN and JAMES T. NELSON, Kent State University

Most learning theories assume that associations gain strength incrementally with repeated pairings of the stimulus and response;[1] recently the incremental assumption has been questioned by Estes and others.[2] In the experiment of interest here, Estes, Hopkins and Crothers, advancing an all-or-none view of learning, presented stimulus-response pairs for a learning trial followed by two tests of recall for the response-member.[3] Finding that responses not recalled on Test 1 were not recalled on Test 2, either, they concluded that these responses had not been learned.

An incremental interpretation of the results of Estes, Hopkins, and Crothers is, however, not completely ruled out. If items are heterogeneous in difficulty, conventional theory can be maintained on the assumption of a threshold of response. Although responses far below threshold would not be expected on either test-trial, responses near threshold might occur often on the second test even if not on the first. This analysis was implemented in the present experiment by two procedural features. First, although Estes has limited his model to recall, a sensitive low-threshold test of recognition was used.[4] Secondly, although half the learning items appeared on the learning list once, half appeared three times. Thus more responses are presumed to lie near threshold than was the case in the earlier study, not only because the threshold is lower but because some items, those presented three times, gain increased strength.

Another feature of the test of recognition must be noted. Estes, Hopkins, and Crothers used an *a priori* guessing estimate of ⅛ for 8 items. Since the test of recognition used here requires a discrimination between

---

* Received for publication December 18, 1961.

[1] See, for example, C. L. Hull, *Principles of Behavior*, 1943, 102-123; K. W. Spence, *Behavior Theory and Conditioning*, 1956, 54-87.

[2] W. K. Estes, Learning theory and the new "mental chemistry," *Psychol. Rev.*, 67, 1960, 207-223; W. K. Estes, B. L. Hopkins, and E. J. Crothers, All-or-none and conservation effects in the learning and retention of paired associates, *J. exp. Psychol.*, 60, 1960, 329-339; Irvin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193; L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, this JOURNAL, 73, 1960, 22-40.

[3] Estes, Hopkins, and Crothers, *op. cit.*, 329-339.

[4] Estes, *op. cit.*, 220-221. Since recognition and recall, however, are commonly regarded as alternative indices of the underlying learning process, differential findings for recognition and recall would be of theoretical interest.

learning-items and control-items, the latter appearing for the first time during test-trials, the frequency of response to control-items can be taken as an empirical rather than *a priori* estimate of the guessing behavior of each *S*.

The test-measure for this study was frequency of response to items on Test 2 not responded to on Test 1. The all-or-none model predicts no difference between learning-items and control-items; incremental theory expects greater values for both once- and thrice-presented items, with the latter value higher.

METHOD

Subjects. The *S*s were 44 volunteers from classes in introductory psychology. Two additional *S*s were excluded from the experiment because of apparatus-failure and *S*s' failure to follow instructions.

Apparatus. Learning-materials were presented to *S* at 4-sec. intervals by means of a Hullian type memory-drum set on a desk. The materials were visible through an opening in a black screen which otherwise obstructed *S*'s view of the apparatus and *E*.

A short lever protruding 3 in. from a wooden box was mounted in front of the screen. *S* responded to the stimulus-materials by pushing the lever 60° to the right of its vertical resting position. A recording apparatus was connected electrically to the lever to provide a record of sweeps of the lever greater than 10°.

Materials. The stimulus-materials were randomly drawn three-letter consonant-syllables. None included the letter *Q*. On the learning-trials, the syllables to be learned were presented along with a large red crayon-dot; these syllables were mixed randomly with fill-in syllables which appeared alone.

The learning-list consisted of six learning-syllables (with red dots) appearing once, and six appearing three times each; these 24 presentations were interspersed with 12 fill-in syllables. The three 'trials' in this learning-list were continuous.

The test-list consisted of the 12 learning-syllables and 6 *new* fill-in syllables, hereafter designated controls, appearing once each for Test 1; immediately following, they appeared in a new random order for Test 2.

Procedure. Each *S* was given a practice-session with five learning-syllables and fill-in syllables. *S* was then given learning-list instructions to the effect that (a) some syllables would have red dots next to them, (b) his task was to remember the red-dot syllables and push the response-lever each time one appeared, and (c) later, his task would be to recognize them without the red dots. The learning-list was then presented.

For the test-trials, *S* was instructed that (a) there would be no red dots, (b) his task was to push the lever each time he thought he recognized a red-dot syllable from the learning-list, and (c) it was permissible to guess. The test-list was then presented. Following this, the procedure was repeated with a new list of syllables. Thus each *S* learned two lists.

The 12 syllables (six from each list) chosen for one learning-list presentation to half the *S*s were used as control-items for the other half.

RESULTS AND DISCUSSION

Table I shows proportions of responses over *S*s and syllables for Test 2. The proportions are responses following responses (p_{RR}) or responses following no responses (p_{NR}) for learning (L) and control (X) items.

The data for items responded to on the second test only (p_{NR}) are of principal interest. The important comparison is between the p_{NR} proportions for controls versus the once- and thrice-presented learning-items.

TABLE I

PROPORTIONS OF RESPONSE ON TEST 2 GIVEN RESPONSE (RR) OR NO RESPONSE (NR) ON TEST 1

Measure	One presentation (L_1)	Three presentations (L_3)	Controls (X)
p_{RR}	0.735	0.833	0.571
p_{NR}	0.288	0.437	0.165

According to an all-or-none hypothesis, the L_1 and L_3 proportions should not be greater than X; otherwise the evidence favors the incremental learning hypothesis. The learning-item proportions, 0.288 and 0.437, exceeded that for controls, 0.165. A test for differences between mean proportions (arcsin transformation) was significant in an analysis of variance ($F = 24.03$; $df = 2/86$; $p < 0.0005$).

Each mean proportion for a learning-item was significantly higher ($p < 0.001$) than the proportion for controls by Scheffé's test for multiple comparisons.[5] Apparently, effects of learning not demonstrated on Test 1 can be shown on Test 2, as would be expected according to an incremental assumption.

The interpretation in accord with incremental theory is further supported by the difference between mean proportions for once- and thrice-presented learning-items. As suggested earlier, a difference in favor of the L_3 items is expected by incremental theory; the observed difference is as expected and significant ($p < 0.001$).

Although irrelevant to the controlled comparison between learning-items and control-items, the question of learning during test-trials deserves some interest. There is little reason to believe that absolutely no learning occurs on the test-trials, despite the absence of experimental feedback or reinforcement. Note the proportion of responses following responses to control-items (p_{RR}). If no learning occurred on Test 1, p_{RR} for these items should be no greater than 0.049, the product of the

---

[5] H. Scheffé, *The Analysis of Variance*, 1959, 66-67.

(independent) proportions of control-responses for each test trial. The observed value was 0.571.

Perhaps the major finding of this study is that when the learning of specific materials by specific individuals is observed, contrary to the ordinary procedures of estimating individual learning from the behavior of a group, there is still little basis for choosing an all-or-none interpretation. All-or-none learning may truly be the case at some more fundamental level of behavior, as all-or-none firing is the case in neural action, but until techniques for measuring behavior at that level are known, an incremental interpretation of the learning process seems more fruitful.

SUMMARY

A design advanced by Estes for testing the assumption of all-or-none learning of paired associates was extended with 44 *S*s to introduce a measure of recognition and repeated presentations in learning. Each of six once-presented and six thrice-presented stimulus-items on a learning-list was paired with a single motor-response, followed by two trials on a test of recognition. The relative frequency of responses on Test 2 following no response on Test 1 for once- and thrice-presented items was 0.288 and 0.437, respectively, as against 0.165 for control-items. On the basis of the hypothesis of all-or-none learning, no differences would be expected. The results were taken to support the view that learning occurs incrementally.

A TEST OF THE RELATIVE-VALENCE THEORY OF INSTRUMENTAL-REWARD LEARNING

By MARCIA J. HEATHCOTE and R. A. CHAMPION, University of Sydney

The common experimental situation of a rat learning to run from a starting box through a straight runway to a goal-box appears to allow some meaningful comparisons of S–R and cognitive interpretations of behavior. According to Spence's hypothesis in S–R terms,[1] a primary variable in this situation is the vigor of the fractional goal-response (r_g) elicited in the starting box and runway (*Rx*) through stimulus-generalization from the goal-box (*Rx*), whereas a cognitive theory such as Tolman's emphasizes the importance of the valence of the goal-box.[2] To add to the deductive power of Tolman's theory and to bring it in line with their data, Gonzalez and Diamond modified it by supposing that the operative factor is the valence of the goal-box relative to that of the starting box and runway, rather than the absolute valence of the goal-box.[3] These experimenters (*E*s) gave rats unrewarded trials in the whole alley as a pretest and then directly placed and fed them in boxes which varied in degree of similarity to the starting box and runway, and the goal-box. Their chief finding was that with alley *RxBy* (goal-box dissimilar to starting box and runway), rats fed in a box (*By*) similar to the goal-box performed better in a later test than did a group fed in a box (*By*) similar to the rest of the apparatus. It is open to the S–R theorist, however, to argue that the rats learned to 'look for food' in the boxes during the direct-placement feedings and that this competing response generalized more to *Rx* from *Bx* than from *By*.

Spence's hypothesis about the role of r_g was originally devised to explain the effects of incentive in instrumental-reward learning and it presumably applies unequivocally to the usual training conditions, as distinct from the less usual, direct-placement conditions. Thus, if a crucial factor is the degree of generalization of r_g from goal-box to starting box and runway, then a group trained with *RxBx* should show superior performance to one trained with *RxBy*. There may be some transfer of the competing, 'looking-for-food response' in the former group, but this response presumably soon

---

Received for publication January 20, 1962.

[1] K. W. Spence, *Behavior Theory and Conditioning*, 1956, 134-164.

[2] E. C. Tolman, Principles of purposive behavior, in Sigmund Koch (ed.), *Psychology: A Study of a Science*, 2, 1959, 92-157.

[3] R. C. Gonzalez and Leonard Diamond, A test of Spence's theory of incentive-motivation, this JOURNAL, 73, 1960, 396-403.

becomes subordinate to the reinforced instrumental response of running. In view of the possibility of such competition, however, a clearer test should be provided if all *S*s were trained with *RxBx* and then tested with *RyBx*, with the similarily of *Ry* to *Bx* systematically varied. If the initial training is continued to a point where competing responses are clearly not occurring, then any interference by the 'looking-for-food' response can be discounted. Both the S–R and cognitive theories should lead to the prediction of a loss of performance immediately after the shift from *Rx* to *By* with subsequent recovery. But whereas the r_g-hypothesis implies that later performance in *Ry* will be directly dependent on the similarity of *Ry* to *Bx*, being superior with greater similarity, the relative-valence hypothesis leads to the opposite prediction, namely that later performance in *Ry* will vary inversely with the degree of similarity of *Ry* to *Bx*, for the valence of *Bx* is relatively greater when *Ry* is dissimilar to *Bx*. The following study was therefore conducted along these lines.

Method: (1) Subjects. The *S*s were 13 experimentally naïve female albino rats, 243 days old at the beginning of the experiment.

(2) Apparatus. The walls of the alley were ¾-in. pine; the lid, clear perspex; and the floors plain aluminum. The internal dimensions of the apparatus being as follows: starting boxes: 5 in. wide, 6½ in. long, and 5½ in. high; runways: 5 in. × 2 ft. × 5½ in.; goal-boxes: 5 × 9½ × 5½ in. One alley was painted white throughout and one black, while dark- and light-gray starting box-runway combinations were used in test-trials; the dark gray corresponded to Value 4 on the Munsell system, and the light gray matched Value 8. The starting boxes were separated from the runways by guillotine-doors of clear perspex which activated a microswitch when raised, so as to start an electronic chronoscope, and the passage of the rat's body across a gap in the metal floor at this point activated a relay which stopped the chronoscope, thus giving a measure of starting-time. A curtain of the same color as the runway on the runway-side separated the runway from the goal-box; the curtains were the same color as the goal-box on the other side. Running times were measured with a stopwatch.

(3) Procedure. The experiment consisted of preliminary training, training, and test. Seven *S*s were trained with an all-black alley *(R_xB_x)* and 6 with an all-white alley *(R_xB_x)*, as a control for the stimulus-dynamism effect. In the test-stage these two groups were subdivided and run with dark-gray or light-gray starting box and runway *(R_y)*. The design of the experiment and the alternative predictions are summarized in Table I. As set out therein, the *S*s tested with high similarity of starting box and runway in training and test (black and dark gray, and white and light gray) are designated as Group A, whereas, *S*s tested with low similarity (black and light gray, and white and dark gray) are designated as Group B.

(4) Preliminary training. The *S*s were allowed to explore the alleys of the training stage in pairs for 30 min. on each of 5 days. There were no doors or curtains in the alleys, and the *S*s were fed *ad lib* in their home cages.

(5) Training. On each of the first 7 days of training the *S*s were given two re-

inforced trials, with an intertrial-interval of approximately 10 min., the reward being a 1-gm. (dry weight) moistened food-pellet in the goal box. A 24-hr. feeding schedule was introduced at this stage, each *S* being allowed to eat moistened food for 1 hr. after training. On each of the following 10 days the *Ss* were given 3 trials, a non-reinforced trial being interpolated between two reinforced trials. On non-reinforced trials each *S* was left in the goal-box for 2 min., the average time taken to eat a 1-gm. pellet.

TABLE I

EXPERIMENTAL DESIGN AND ALTERNATIVE THEORETICAL PREDICTIONS

Training conditions	Test-conditions	Predictions	
		S–R	Cognitive
All-black alley $(Rx\ Bx)$ $N=7$	(A) $Ry=$dark gray $N=4$	A>B	B>A
	(B) $Ry=$light gray $N=3$		
All-white alley $(Rx\ Bx)$ $N=6$	(A) $Ry=$light gray $N=3$	A>B	B>A
	(B) $Ry=$dark gray $N=3$		

(6) Test-stage. On each of the next 8 days, the *Ss* were given 2 trials; the first was a reinforced trial in the training apparatus and the second was a test-trial, consisting of a non-reinforced run with the dark- or light-gray starting box and runway.

Results. The results of the test with running time as the measure are shown in Fig. 1, the long and variable times for the first test-trial being omitted, and it may be observed that Group B is superior to Group A. The two groups were strictly comparable late in training before any shift from Rx to Ry, both giving a mean running time of 0.69 sec. over the last six training trials. An analysis of the variance in the mean running times over the last seven test-trials, on the other hand, showed the difference represented in Fig. 1 to be statistically significant ($F = 6.82$, $p < 0.05$ for 1/10 *df*). The *Ss* trained with the white alley showed a superior performance to those trained with the black alley, and this difference persisted on test-trials ($F = 18.18$, $p < 0.01$ for 1/10 *df*), but there was no interaction between this variable and the factor distinguishing Groups A and B. The results with starting time as a measure were qualitatively similar to the data on running time but their greater variability precluded statistically significant differences.

The results of this experiment are in keeping with the relative-valence hypothesis and run counter to the expectation derived from the assumed mode of operation of the r_g-mechanism. Whereas any dynamogenic effect of r_g on the instrumental act may have been masked under other conditions by the activity of 'looking for food,' it is unlikely that any such interference could have occurred in the present study in view of the use of reinforced training trials and the absence of direct-placement feedings.

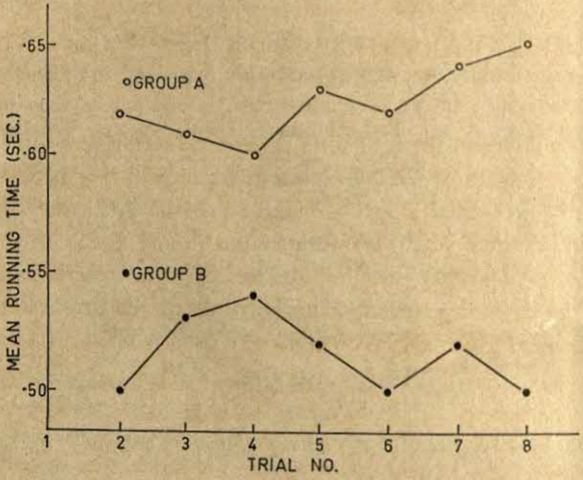


FIG. 1. PERFORMANCE ON THE LAST SEVEN TEST-TRIALS

SUMMARY

After training in an all-black or all-white alley, some members of a group of 13 rats were given test-trials with the starting-box and runway changed to dark gray or light gray respectively (Group A, $N = 7$), whereas the remaining rats were tested with light gray or dark gray starting-box and runway (Group B, $N = 6$). It was expected that all the *Ss* would show an initial loss of performance with the shift, but whereas Spence's r_g-hypothesis implies that later performance should be superior in Group A (starting-box and runway more similar to goal-box), the relative-valence hypothesis of Gonzalez and Diamond leads to the predicted superiority of Group B, the valence of the goal-box being relatively greater with starting-box and runway less similar to the goal-box. The results supported the relative-valence hypothesis.

INTENTIONAL AND INCIDENTAL LEARNING UNDER DISTRACTION

By WILMA A. WINNICK and RICHARD A. LERNER, Queens College, New York

Studies of incidental learning—its general characteristics in comparison with intentional learning and the parameters of which it is a function—have generally proceeded along two lines: (1) the incidental subject (*S*) serves as the experimenter (*E*) for the intentional *S* and in this role is presumed to have no set or intention to learn the material; and (2) a single *S* is given both relevant material to learn and irrelevant material about which he is given no instructions and usually no information.

The first procedure, in which incidental and intentional learning are both directed toward the same material, has been used most fruitfully by Postman and his colleagues.[1] They have shown, for example, among other contrasts with intentional learning, that incidental learning is less sensitive to serial order or linkage among items, and less affected by length of list and less subject to intra-serial interference.

The second procedure, in which materials differ but the *S*s are the same for the two kinds of learning, has been employed in studies interested in interactions between the two forms of learning as they occur simultaneously. Thus, Bahrick found that if the incentive for intentional learning is increased, intentional learning will speed up, but the amount of incidental learning is decreased.[2] On the other hand, Kausler, Trapp, and Brewer found, when an emotionally based drive was employed, that both incidental and intentional learning were increased.[3] The contrast in findings of these two studies indicates the difficulty of predicting the effect of a variable upon incidental learning, even though its effect upon intentional learning is known.

---

* Received for publication August 1, 1962. This study was supported by an Undergraduate Research Grant (No. 2M-7311) of the National Institute of Health.

[1] Leo Postman and L. W. Phillips, Studies in incidental learning: I. The effects of crowding and isolation, *J. exp. Psychol.*, 48, 1954, 48-56; Leo Postman, P. A. Adams, and L. W. Phillips, II. The effects of associative value and the method of testing, *ibid.*, 49, 1955, 1-10; III. Interserial interference, *ibid.*, 51, 1956, 323-328; IV. The interaction of orienting tasks and stimulus-materials, *ibid.*, 51, 1956, 329-333; Leo Postman, P. A. Adams, and A. M. Bohm, V. Recall for order and associative clustering, *ibid.*, 51, 1956, 334-342; Postman and Adams, VI. Intraserial interference, *ibid.*, 54, 1957, 153-168; VII. Effects of frequency of exercise and length of list, *ibid.*, 56, 1958, 86-95; VIII. The effects of contextual determination, *ibid.*, 59, 1960, 153-164; Postman and Phillips, IX. A comparison of methods of successive and single recalls, *ibid.*, 61, 1961, 236-241.

[2] H. P. Bahrick, Incidental learning under two incentive conditions, *J. exp. Psychol.* 47, 1954, 170-172.

[3] D. H. Kausler, E. P. Trapp, and C. L. Brewer, International and incidental learning under high and low emotional drive levels, *ibid.*, 58, 1959, 452-455.

The present experiment was directed toward a different, but possibly related, interaction between intentional and incidental learning—the effect of distraction. A number of studies have been concerned with the effect of various distractors on intentional learning.[4] The second procedure mentioned above makes possible an examination of the effects of distractions acting simultaneously upon intentional and incidental learning. It was expected that differences in the effect of distractors upon intentional and incidental learning might reveal differences between the kind of learning which occurs as a consequence of instructions from *E* and the learning which occurs under self-instructions or as a transfer from other experiments on learning.

Method: (1) Subjects. The *S*s were selected from introductory classes in psychology at the Queens College summer session and were assigned randomly to four groups of 20 *S*s each.

(2) Apparatus. On a Lafayette memory-drum, a list of 12 girls' names was exposed at a 2-sec. rate with a 32-sec. interval between successive trials. Each name was typewritten in a 2¼ by ¼-in. rectangle of color, ensuring that there was spatial and temporal proximity between each name and its associated color. There were 6 colors and 12 names; every color being paired with two of the names, as follows: Muriel, gray; Vivien, yellow; Abigail, red; Georgia, brown; Bridget, green; Jessica, gray; Ingrid, blue; Natalie, red; Rosaline, yellow; Leonora, brown; Claudia, blue; and Eleanor, green. For the three distractions, the following were played continuously by a tape recorder: (a) a list of 135 girls' names, all differing from those to be memorized; (b) a simultaneous dual recording of Arnold Shonberg's atonal "Pierrot Lunaire" and Tito Puente's "Top Percussion"; and (c) selections from various radio programs, including news broadcasts, music, discussions, and commercials.

(3) Experimental procedure. All the *S*s were given standard instructions to familiarize them with the method of anticipation and were told to memorize the list of names by this method. Colors were not mentioned in the instructions. One of the groups of *S*s served as a control group, learning the names in a quiet situation, while each of the other three groups learned the names while listening to one of the three recordings.

Training continued to the criterion of one errorless trial. When this had been reached, the *S*s were immediately shown a chart with the six colors in ¾ by ½ in. rectangles across the top and with the names in order down the side. Each *S* was told to match each name with its associated color; the number of correct matches being his score in incidental learning.

Results and discussion. As Table I reveals, each distraction-group showed, in comparison to the control group, a loss of efficiency of intentional learning. Analysis of variance revealed an *F* for the inter-group difference of 4.01, close to the *F* of 4.04 required for significance at the 1% level.

---

[4] Hulsey Cason, The influence of attitude and distraction, *ibid.*, 22, 1938, 523-546.

Tukey's gap-test revealed significance among the three comparisons of the experimental-control groups, but none among the experimental groups. Each distraction used effected a change in efficiency of intentional learning, but none more so than either of the others.

Although one might predict that the greater number of trials taken by the distracted groups might favor higher incidental scores for them, their scores were found to be slightly, but not significantly, lower than those for the control group. Analysis of covariance was used to test for a possible difference when incidental scores were corrected for initial differences in number of learning trials. This analysis showed that the residual difference lacked any significance (corrected $F = 0.075$). The distractions, therefore, had little effect on incidental learning in comparison to their effect on intentional learning. The extra trials taken by the groups learning under distraction were not apparently of aid in furthering incidental learning.

TABLE I

TRIALS TO LEARN INTENTIONAL MATERIAL AND INCIDENTAL SCORES (COLOR) FOR THE CONTROL AND THE THREE EXPERIMENTAL GROUPS

Groups of *S*s	Trials to learn		Number of colors learned	
	Mean	*SD*	Mean	*SD*
Control	12.20	3.52	3.70	2.88
Recorded names (a)	16.75	5.15	3.35	2.26
Atonal music (b)	18.95	7.63	3.25	2.14
Radio (c)	17.30	7.80	3.40	2.10

Comparisons of the *SD*s among the groups revealed larger differences in variability for the distracted groups in intentional learning than for the control group. *F*-comparisons were made of *SD*s of the separate distraction-groups and of the control group, with the following outcome: Distraction-Group (a), $F = 2.14$, significant at 5% level; Distraction-Group (b), $F = 4.70$, significant at 1%, and Distraction-Group (c), $F = 4.91$, significant at 1% level. Distractive effects produced an increase in inter-*S* differences in intentional learning, without comparable effect upon incidental scores, which showed instead a small, insignificant decrease in variability.

The contrast in the effect of the distractions upon these two forms of learning may be related to the lack of susceptibility to sources of intralist interference reported in a study of Postman and Adams.[5] In that study, incidental learning was shown to suffer less of an adverse effect from increases in length of list, and to enjoy less benefit from added repetitions than did intentional learning. Furthermore a previous study had shown incidental learning to be less subject to intra-serial interference.[6]

The immunity of incidental learning to interference outside the learning situation, as well as to interferences within the task itself, seems indicative of the strength of preëxperimental learning habits, possibly generated under varying, sometimes distracting, environmental conditions. The kind of learning which *S* does under direction from *E*, on the other hand, might be viewed as requiring effort on *S*'s part and is, therefore, vulnerable when *S* must exert effort to overcome distractors.

---

[5] Postman and Adams, VII. Effects of frequency of exercise and length of list, *J. exp. Psychol.*, 56, 1958, 86-95.

[6] Postman and Adams, VI. Intraserial interference, *ibid.*, 54, 1957, 153-168.

APPARATUS

THE ELECTRICAL QUANTIFICATION OF THE CONSUMPTION OF LIQUIDS BY RATS

By WILLIAM W. DAWSON, RONALD S. LINDNER, and NATHAN PERRY,
Florida State University

Methods for quantifying the consumption of liquids may be divided into three types. One is that of Hill and Stellar and T. A. Hunter, which used the animal to complete an electrical circuit during drinking.[1] A second type, described by Duffy and Price, is that in which liquid level-changes are recorded by direct measurements of the change of the height of the miniscus.[2] The third type, perhaps the most common, requires a learned manipulation by the animal which may be rewarded by the presentation of the liquid.[3] When changes of drinking behavior *per se* are of interest, as in the studies of Garcia and Kimeldorf *et al.* or Young and Falk, it is desirable to maintain a high degree of similarity between the normal drinking activity and that required to obtain the experimental fluid.[4] The method described below offers advantages in both quantification and continuity of the environment.

By this method the number of licks at a standard drinking tube is recorded in the home cage. As has been reported, rats lick at a stable rate and consume equal amounts of fluid at every lick.[5] It may be assumed, therefore, that licking is a valid indicator of liquid-consumption.

The circuit presented in Fig. 1 is similar in some respects to early vacuum tube circuits of Hill and Stellar and of Hunter.[6] It is believed, however, that the gas tube lends many advantages as an 'all or none' actuator. Previous experience with vacuum tube relay-circuits has yielded evidence of sluggish and variable operation. Elements C_1 and R_1 of the input circuit largely eliminate 60-c.p.sec. interference. Previously these transients have been the major factors which encouraged the use of mechanical devices in which multiple installation and long transmission lines were necessary. Simply, this circuit is a high input impedance, gas tube switch. A voltage increase at the grid of V_{1a} produced by a lick at the fluid column causes an increase in the conduction of V_{1a} and an elevation of its cathode voltage. The cathode can furnish both modest current, and voltage to the grid of V_2. R_2 is so adjusted that, at rest V_2 is held in a non-conducting state by a negative 3-v. bias. The increase in the potential at the vacuum tube cathode can then overcome this bias so that threshold for V_2 is reached and the tube conducts completely. Conduction of V_2 actuates the relay and the associated recording device. By this sequence each lick of the animal is recorded.

---

[1] J. H. Hill and Eliot Stellar, An electronic drinkometer, *Science*, 114, 1951, 43 f.; T. A. Hunter, An electronic contact relay, this JOURNAL, 65, 1952, 458-459.

[2] M. L. Duffy and G. E. Price, A continuous drinking recorder for small animals, this JOURNAL, 69, 1956, 664 ff.

[3] B. F. Skinner, Thirst as an arbitrary drive, *J. gen. Psychol.*, 15, 1936, 205 ff.

[4] J. Garcia, D. J. Kimeldorf, E. L. Hunt, and B. P. Davis, Food and water consumption of rats during exposure to gamma radiation, *Rad. Res.*, 4, 1956, 33 ff.; P. T. Young and John L. Falk, The relative acceptability of sodium chloride as a function of concentration and water need, *J. comp. physiol. Psychol.*, 49, 1956, 569-575.

[5] Hill and Stellar, *op. cit.*, 43; P. T. Young and W. F. Kappauf, Apparatus and procedures for studying taste-preferences in the white rat, this JOURNAL, 75, 1962, 482 f.

[6] Hill and Stellar, *op. cit.*, 43; Hunter, *op. cit.*, 458 f.

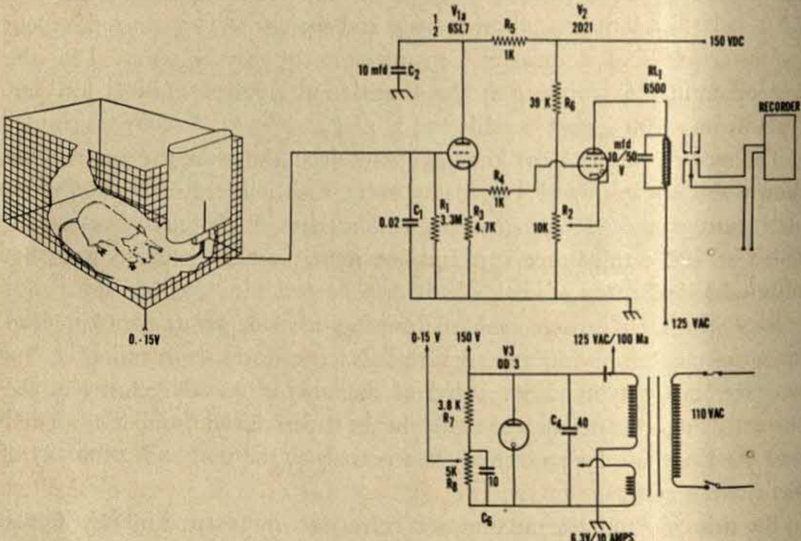


FIG. 1. DIAGRAM OF THE SENSING AND POWER SUPPLY UNITS

All capacitance values are microfadards with 150-v. DC ratings, except where noted. Voltage-readings were taken with a 20,000 Ω per volt meter. Resistance R_2 and R_5 are rated at 2 watts. All others are 1 watt.

As long as the animal remains in contact with the fluid, Column V_2 remains in the conducting condition. Once the fluid contact is broken, V_2 becomes non-conducting within 17 m.sec. It will not conduct again until a new connection initiates the firing sequence. Adjustment of the grid bias at R_2 also acts as a sensitivity-control. Over-adjustment of this control can result in an above-threshold bias-condition in which RL_1 will remain actuated and insensitive to licking.

Since V_{1a} is a vacuum tube, the current through the animal which is required to operate its grid can approach zero. For stable operation, however, $C_1$and R_1 are required. Voltages up to $+10$ can be applied to the cage with currents remaining below 2.5×10^{-6} amp. This level is below the 7.7×10^{-5} threshold for electro shock and the 1.0×10^{-3} amp. threshold for taste threshold stimulation.[7] In operation R_8 may be adjusted for the best operating bias or so that the animal circuit current flow is just perceptable on a 100-μ amp meter.

The power supply is of the half-wave type and with the components shown will operate up to 10 sensing units. Each unit requires 7.5-m. amp. plate current and 750-m. amp. filament current. Additional units may be added if the transformer and rectifier ratings are adjusted to meet the additional load.

Rats which have been subjected to long-term deprivation often grasp the drinking tube with their teeth. This causes erratic operation of the recorder, but may be disregarded as the animal quickly returns to the normal drinking rate. It frequently helps if the drinking tube is inserted into the cage just far enough to be accessible to the tongue but not to the teeth. Some stoppage of liquid or spillage may occur if air bubbles collect in the drinking tube. A piece of non-corrosive wire inserted into the bottle end of the tube and extending almost to the nipple will usually prevent the collecting of bubbles.

Some consideration must be given to the fluid itself. Voltage-changes are conducted through the fluid to operate the grid of V_{1a}. This presupposes that the fluid contains an ion content. High quality distilled water has no ion content and cannot be used with this sensing device. Tap water with its high ionic concentration is satisfactory. A more consistent fluid can, however, be made from double distilled water and NaCl added to produce a concentration of less than 0.009% by weight which has been established as threshold for the rat.[8] Concentrations lower than 0.001% by weight have been found to provide sufficient conduction of this sensing section.

---

[7] K. F. Muenzinger and R. H. Mize, The sensitivity of the white rat to electric shock: Threshold and skin resistance, *J. comp. Psychol.*, 15, 1933, 148 f.

[8] W. J. Carr, The effect of adrenalectomy upon the NaCl taste threshold in rat *J. comp. physiol. Psychol.*, 45, 1952, 379.

Use of this method continuously for 18 mo. discloses the following advantages over the earlier techniques.

(1) The circuit is fast. Single licks may be totalized over long periods.

(2) Many of the various recording systems may be actuated by this sensing system.

(3) No alteration of the animals' consummatory habits or environment is required.

(4) The system is silent and all secondary cues can be eliminated by isolation of the recording system.

(5) Nocturnal patterns are not disturbed by lights for photo-cells or liquid level-readings.

(6) The system will provide consistently stable performance when subjected to transients, long input cables, and moderate variations of line voltage.

A HIGH-SPEED CONSTANT-ILLUMINATION TACHISTOSCOPE

By HERBERT ZIMMER, C. L. BOLEN, and B. D. NICORA,
University of Georgia

The circuit described here shows the power supply, switching, and driver units of a constant-illumination tachistoscope, which was designed to present a long series of visual stimuli for accurately timed durations.[1] It can as readily be used as a flicker-fusion apparatus with accurately controlled square-wave light pulses of continuously variable durations and off-intervals, by providing the appropriate setting for the light. The circuit governs the operation of two modified R1131C glow modulator tubes, each mounted in a separate and identical automatic slide projector, one providing background illumination, the other carrying the stimulus-slides. Some projectors have p rovision for hooking successive slide trays together, making an uninterrupted presentation of *n* slides feasible. Since the light of the R1131C is concentrated in a narrow beam, the space and optical characteristics of some projectors may necessitate a broadening of the lightbeam by enlarging the hole in the two mica disks which face the top end of the tube.

The onset of a single, 18-v., square-wave trigger pulse shuts off the background light and turns on the stimulus-lamp simultaneously. The termination of this pulse effects an instantaneous reversal of this condition. Fig. 1 illustrates the time relationships. The trigger pulse may be available in existing laboratory equipment, or it may be obtained from a timer, counter, or wave form generator, preferably one with at least three pre-sets. One pre-set is required to time the onset of the stimulus, a second for its termination, and a third to advance the projector to the next stimulus-slide. Additional work can be performed by these pulses. In the present application, the stimulus-onset pulse also starts a reaction-timer, and the slide-changing pulse resets it after it has been read.

The speed of this tachistoscope is accurate to better than 1 m.sec., provided that the trigger pulse has a rise-time in the microsecond range. A number of wave form generators are available to supply pulses of microsecond rise-time. While the rise-time of an incandescent lamp is 50-200 m.sec., and its decay-time somewhat longer, the rise-time to maximal brightness of a glow modulator tube is about 0.5 m.sec., and its decay-time

---

[1] This study was supported by the Air Force Office of Scientific Research Grant AF-AFOSR-257-63.

of like magnitude. Therefore, the shortest dependable 'on-time' at maximal brightness is about 1 m.sec. If maximal brightness is not required, the 'on-duration' can be decreased to as little as 5 μsec., with correspondingly less light output. The R1131C is the slowest component of the circuit. Switching time of the rest of the circuit is estimated at a fraction of a microsecond. For this reason it does not limit the performance of the glow modulator tubes.

The wave length of most of the light generated by the R1131C glow modulator tube lies between 590 and 650 mμ.

An illuminometer is required to equate the brightness of stimulus- and background-lamps. The brightness of each lamp can be adjusted manually and individually by means of the two 20K potentiometers. Within the operating range of the tachistoscope driver, 0-30 m.amp. the relation between R1131C tube current and light output is nearly linear. Tube life can be prolonged by operating at the lowest amperage which gives sufficient light for any particular application.

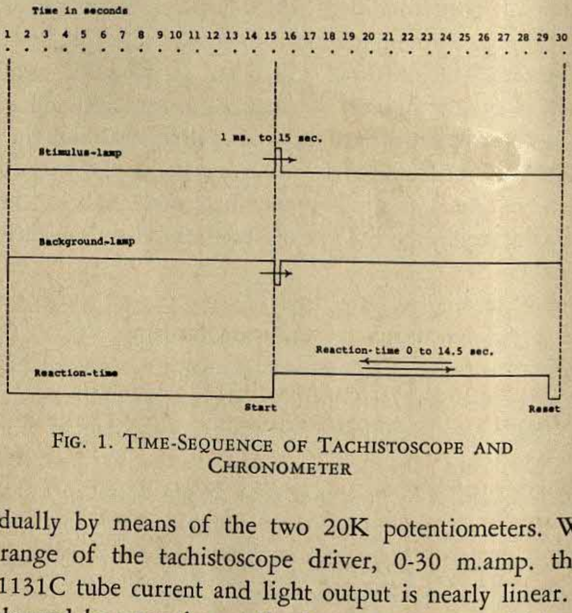


FIG. 1. TIME-SEQUENCE OF TACHISTOSCOPE AND CHRONOMETER

The power supply employs one 5Y3 dual diode, functioning as a full-wave rectifier, and one 6X4 dual diode, operating as a half-wave rectifier, to produce +350 and −355 v. (See Fig. 2.)

The electronic switch consists of one 12AU7 dual triode and associated resistors. The plate of one triode is tied to the grid of the other triode, permitting only one of the triodes to conduct at any given time.

Each of the two R1131C glow modulator tubes is driven by a 6AQ5 beam-power amplifier tube, which is in turn controlled by one half of the 12AU7 dual triode.

While the normally-on background lamp is burning, its 6AQ5 is conducting, but its triode is not. A square-wave pulse, with a minimal amplitude of +18 v., applied to the grid of the background triode, produces the following changes: (1) The background triode conducts. (2) The background 6AQ5 turns off, as a negative voltage is applied to its grid. (3) The background R1131C lamp goes off, as soon as its 6AQ5 driver does. (4) The stimulus-triode turns off, as its grid becomes negatively charged. (5) The normally-off stimulus 6AQ5 conducts, when its triode is shut off and its grid goes positive. (6) The stimulus R1131C lamp goes on, as soon as its 6AQ5 driver is activated.

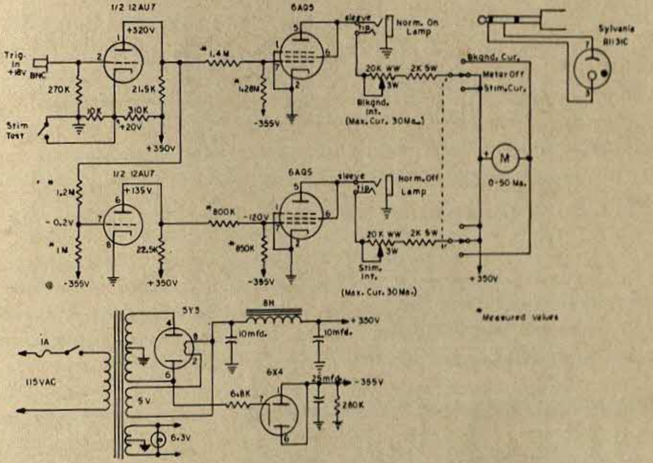


FIG. 2. SCHEMATIC DIAGRAM OF TACHISTOSCOPE POWER-SUPPLY, SWITCHING, AND DRIVER CIRCUITS

Several years of operation have shown this simple, inexpensive circuit to be effective and dependable, requiring a minimum of maintenance or adjustment.

THE MULTIPURPOSE PSYCHOLOGICAL LABORATORY AT AMHERST COLLEGE

By JOHN W. DAVENPORT, Amherst College

In 1961, a lecture hall at Amherst College was converted into a 17-room psychology laboratory designed for undergraduate laboratory exercises, staff research, and various other uses.[1] Foremost among the considerations in its design was our desire to modernize the weekly laboratory exercises of the Introductory Course, and at the same time create a new type of facility for testing out ideas relating to the logistics of gathering data in common areas of psychological research. For the floor-plan see Fig. 1.

FEATURES AND USES

Room arrangement. Sixteen small rooms, called 'Remote Rooms' (*RR*s), are connected by cables to a 'Master Control Room' (*MCR*). For control purposes, the *RR*s are completely enclosed, nearly identical in size (approx. 8 ft. sq.), and have filled-cinder-block walls which are relatively sound-resistant. Open-ended vertical conduits, embedded into each room's walls, extend into a 42-ft. sq. attic located above the *RR*s. While mainly designed for convenience in interconnecting rooms, the attic also has potentialities as an area for concealed observation of *S*s in the *RR*s below.

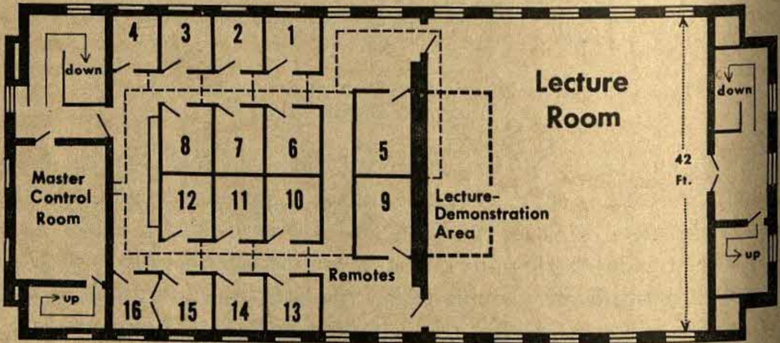


FIG. 1. FLOOR-PLAN OF THE LABORATORY IN RELATION TO ADJACENT SPACE FOR LECTURES, BRIEFINGS, DISCUSSIONS, AND DEMONSTRATIONS

Cable systems. The network of cabling which interlinks the rooms includes multi-conductor cables (36 low-voltage conductors between the *MCR* and each *RR*), coaxial cables (for communication systems), and ordinary AC and variable AC-DC power lines (which originate at control panels in the *MRC*). The multiconductor cables are terminated by Cinch-Jones sockets and Nu-Way stud-panels in each *RR* and by a 70-socket patch panel in the *MCR*. By plugging short multiconductor jumpers into this patch panel, one can connect: (a) any *RR* to any other *RR*; (b) several control apparatuses in the *MCR* to several different *RR*s; or (c) a single control setup in the *MCR* to all *RR*s simultaneously.

---

[1] The laboratory was designed by Professors Theodore Koester, Robert C. Birney, Robert F. Grose, Mr. Robert Heidrich, the College Engineer, and the author.

MCR control, recording, and event-monitoring equipment. The master-remotes principle is applied to the major equipment systems as well as to the laboratory's floor plan. For controlling stimulus-events in *RR*s, the *MCR* is equipped with 'master programming racks' which plug into the multiconductor cable system's patch panel and house an abundance of timers, steppers, relays, etc., of the type familiar in operant conditioning research. Conventional polygraphs can be connected to the apparatus in the *RR*s in a variety of locations within the *MCR*, and several indicator-lamp panels are also provided for the visual monitoring of both stimulus- and response-events in *RR*s and for miscellaneous signaling purposes. The most important of these panels is the 'Toteboard,' which is a rectangular matrix of 132 Dialco Datalites providing a continuous visual indication of up to 11 events (or scores, ratings, etc.) in each of 12 rooms simultaneously.

Communication systems. The laboratory's equipment also includes closed-circuit television and audio systems. A Sylvania Direct-Wire *TV* camera (usually located in the *MCR*, but easily portable for televising anywhere in the building) is cabled to 17-in. *TV* monitors in 12 of the *RR*s and additional outlets in the lecture hall. This system is of service: (a) as a means of instantaneously duplicating and magnifying verbal and pictorial material constituting stimulus-presentations (*e.g.* nonsense-syllables, *TAT* cards); (b) in televised demonstrations to students in *RR*s or in the lecture hall; (c) for remote observation of *S*s' behavior in research studies; and (d) in conjunction with the Toteboard, as part of an informational feedback. In this last case, simply televising the Toteboard response-indications to all *RR*s at once enables a given *S* in a laboratory exercise, or in a social-influence experiment, to see how all the participating *S*s are responding, such that he can note individual differences and group-data trends at successive stages of the experiment.

Like the television system, the audio-system is designed for specific stimulus-presentation and informational-feedback purposes as well as for general communication. It consists of a master audio-console in the *MCR* cabled to speakers, earphones, microphones, and indicator lamps in each *RR*, and is a versatile system providing for one-way or two-way communication between the *MCR* and any number of *RR*s at a time. Its master console accepts signals from tape-recordings, microphones, oscillators, and motion picture sound tracks.

RR equipment. For many laboratory exercises, and certain research studies, the *TV* monitors and audio-components represent all the equipment that is needed in the *RR*s, while for more 'advanced' research, highly complex and specialized devices may be needed. Between these extremes stands a considerable range of instructional and research situations having many requirements in common; to satisfy these, and to keep week-to-week equipments changes in the *RR*s at a minimum, a number of multipurpose devices were constructed and duplicated for the 12 *RR*s used in the Introductory Course. These consist of specially-designed stimulus-light panels, switch panels, manipulanda, animal test-chambers, timing devices, etc., which can be controlled remotely from the *MCR* or operated manually within each *RR*. Used in con-

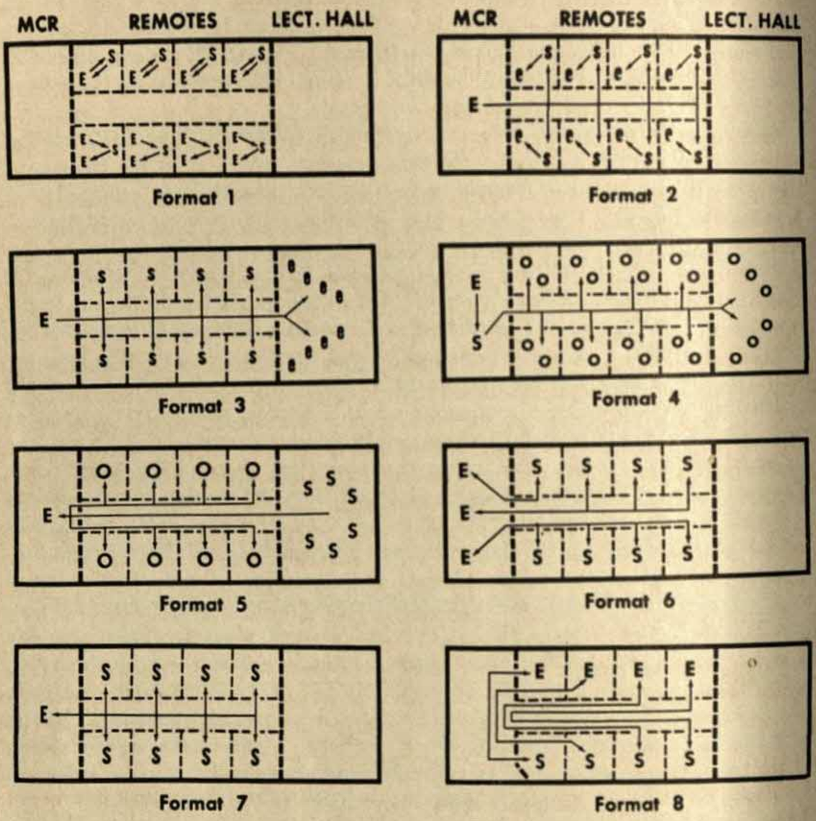


FIG. 2. LOGISTICAL FORMATS FOR LABORATORY EXERCISES, DEMONSTRATIONS, AND RESEARCH STUDIES

Format 1. Traditional Arrangements of *S*s and *E*s in *RR*s only.

Format 2. Synchronized Stimulus Presentations by Master *E* in *MCR;* recordings by sub-*e*s in *RR*s.

Format 3. Same as Format 2, Except That Sub-*e*s Monitor Presentations and *Ss'* Behavior in Lecture Room via *TV*.

Format 4. Behavior of One *S* in *MCR* Observed by *O*s in *RR*s or Lecture Room, via *TV*.

Format 5. Group Behavior Observed by *O*s in *RR*s through Communication with Master *E* in *MCR*, via *TV*.

Format 6. Three Independent Experiments with Varying Numbers of *Ss*, Controlled Simultaneously from the *MCR*.

Format 7. Simultaneous Presentations to and Recording from Many *S*s by a Single *E*.

Format 8. Independent Experiments Using *RR*s as Control Rooms.

junction with the multiconductor and communication systems, these devices play a major role in exercises on reaction-time, probability learning, spatial generalization, verbal learning, and operant conditioning, and serve important switching, signaling, and convenience functions when other research areas are under study.

Logistical formats. The over-all flexibility of the laboratory is indicated in the diagrams of Fig. 2, which represent the more useful of the various spatial-temporal arrangements of *E*s, *S*s, *O*s, and equipment made possible by the several master-remote systems. (As comparison with Fig. 1 will indicate, these diagrams are designed to portray functional relationships, rather than to represent accurately the laboratory's floor-plan.) In terms of the most likely usage, Formats 1-5 are primarily instructional, while Formats 6-8 represent research arrangements. With a little imagination, however, one can conceive of using the first five in serious research (particularly social-interaction research) or the last three as instructional-demonstrational formats.

Especially noteworthy for their pedagogical and experimental advantages over conventional arrangements (Format 1) are the centrally-controlled Formats (2, 3, and 7) involving synchronized stimulus-presentations to many *RR*s at once. In the case of laboratory exercises, these provide important forms of rapid feedback to *S*, both while gathering data (when Toteboard information can be televised to all *RR*s in trial-by-trial fashion) and in class discussions held immediately thereafter (the scheduling of which becomes much easier when collecting data ends simultaneously for all students). In research situations, *E*s employing these formats can enjoy the convenience of having *S*s come to the laboratory in groups at a time, collect a large amount of data in a short time, and can achieve substantially greater control over extraneous variables, especially temporal ones, in the process. These and other advantages were realized in the first year's experience with the laboratory, during which over 40 laboratory exercises, demonstrations, and independent research projects were conducted, and in which more than a dozen research areas of psychology were represented.

The construction of the laboratory's rooms and the installation of permanent heating, ventilation, and electrical systems, which was done under contract, constituted a major expense, but one which we believed was well justified by the obvious experimental-control advantages of enclosed rooms. Most of the laboratory's equipment was constructed in do-it-yourself fashion by four undergraduates working with the author in a 10-week period.[2] This held down equipment costs considerably. Further information about the cost of the laboratory, its physical details, and the equipment-fabrication project is given in a document (No. 7720) which has been deposited with the American Documentation Institute, Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C.

---

[2] Peter L. Rogers, Anthony S. Niskanen, John S. Duryee, and Robert J. Mignone, to whom the Department is deeply indebted.

NOTES AND DISCUSSIONS

THE EFFECTS OF STIMULANT AND DEPRESSANT DRUGS ON PHYSICAL PERSISTENCE

Costello and Eysenck predicted, on the basis of Eysenck's general theory, that pain would be a stronger negative drive for introverts than for extraverts and that this would result in the extravert's physical persistence being greater than that of introverts. Data in line with this prediction were presented.[1] Brachman and Costello, however, in an attempt to replicate these findings, found that the extraverts in their study did not differ significantly from the introverts in physical persistence.[2] They suggested that factors other than physiological ones may have produced the significant findings in Costello and Eysenck's study. It has been found that group conditions of testing, when compared with individual testing conditions, produced an increase in physical persistence.[3] This appears to be in line with the suggestion that factors other than physiological ones are of importance in determining physical persistence.

In the experiments reported in this paper, a test was made of the physiological explanation advanced by Costello and Eysenck to account for the difference between extraverts and introverts. This was done with the use of Eysenck's drug-postulate which states that, "Depressant drugs increase cortical inhibition, decrease cortical excitation, and thereby produce extraverted behavior patterns whereas stimulant drugs decrease cortical inhibition, increase cortical excitation, and thereby produce introverted behavior patterns."[4]

On the basis of Costello and Eysenck's physiological explanation of persistence-differences between extraverts and introverts, and the drug-postulate of Eysenck's general theory, the following predictions were made: (1) Depressant drugs (sodium seconal and meprobamate) will produce an

---

[*] From the Psychological Laboratory of the Regina General Hospital, Regina. This work was supported by a grant from Wallace Laboratories, New Jersey. Mrs. Gloria Glab, Research Assistant, conducted the experiments.

[1] C. G. Costello and H. J. Eysenck, Persistence, personality, and motivation, *Percept. mot. Skills*, 12, 1961, 169-170.

[2] H. M. Brachman and C. G. Costello, The relationship between persistence and personality, *J. Psychol.*, 55, 1963, 299-306.

[3] Brachman and Costello, *op. cit.*, 302.

[4] D. S. Trouton and H. J. Eysenck, The effects of drugs on behaviour, in H. J. Eysenck (ed.), *Handbook of Abnormal Psychology*, 1961, 643.

increase in physical persistence when compared with no-drug or placebo-conditions. (2) A stimulant drug (dextro amphetamine) will produce a decrease in physical persistence when compared with no-drug or placebo-conditions.

(1) In the first experiment, 50 *Ss* were used. There were 25 men and 25 women ranging in age from 17 to 43 yr. They were selected from a larger group of 167 college students (87 men and 80 women) on the basis of their extraversion scores on the Maudsley Personality Inventory.[5] To rule out any possible effects of the extraversion-introversion personality variable, only those *Ss* were used whose extraversion scores were less than 1 *SD* from the mean for the total group. The two groups, one of men and the other of women, were each divided randomly into five groups of five *Ss* each.

Each of the five groups of men and women was allotted randomly to one of five drug conditions—no-drug; placebo; 3 gr. sodium seconal; 10 mg. dextro amphetamine; 800 mg. meprobamate.

S's strength was measured with a hand dynamometer at the beginning of the session. Every *S* was tested individually during one session. His persistence was measured 1.5 hr. after administration of the drug. In the case of the no-drug condition, persistence was measured 1.5 hr. after the arrival of *S* at the laboratory.

Each *S* was given four trials of strength on the dynamometer using alternately right and left hand and beginning with the right hand. There were 1-min. intervals between trials. Two-thirds of the mean strength of each hand was calculated and *S* was given two trials of persistence at this setting. *S* was instructed to keep the pointer as long as possible at this setting for his right hand with his right hand, and at this setting for his left hand with his left hand. The two trials correlated 0.80 and their mean was *S*'s persistence-score. There was a 1-min. interval between trials of persistence.

Analysis of variance of the data revealed that the drug-effect was the only significant effect ($F = 2.650$ with 4 and 44 *df*). Neither sex nor sex interacting with drug-conditions was significant. A test of the difference between the mean persistence for the five drug-conditions indicated that the only significant difference was between the no-drug condition (mean persistence = 19.40 sec.) and dextro amphetamine condition (34.40 sec.), ($t = 2.620$, df 44, $p < 0.02$).

(2) In the second experiment, the *Ss* were 80 girls selected from a group of 164 student nurses. Only those *Ss* were selected whose extraversion-score on the Maudsley Personality Inventory fell within 1 *SD* from the mean for the group. The *Ss* were split randomly into five groups of 16 *Ss* each and each group received one of the five drug-conditions listed for the first experiment. The same procedure was used as in the first experiment.

---

[5] H. J. Eysenck, *The Maudsley Personality Inventory*, 1959.

The only significant differences obtained by *t*-tests were between the mean for the dextro amphetamine condition and each of the means for the other four drug-conditions—the mean for the dextro amphetamine condition being the larger mean in every case. The *t*-values ranged between 3.657 to 7.713 which with 32 *df* indicated significance at the 0.1% level.

The results of both experiments are contrary to prediction. They cast doubt on the physiological explanation proposed by Costello and Eysenck for the greater persistence of extraverts than introverts in their study.

Saskatchewan, Canada C. G. Costello

TABLES VS. GRAPHS AS MEANS OF PRESENTING EXPERIMENTAL RESULTS

The purpose of this Note is twofold: First, and incidentally, to correct an error in a line-cut, for which I must be held responsible, that was made in this Journal many years ago; and secondly, and principally, to deduce from that error some evidence regarding the relative value of tables and graphs in the presentation of experimental results.

Some years ago, when composition and engraving were relatively inexpensive, authors used both tables and line-cuts in their reports. Now, however, because of expense, the number of articles employing both is greatly reduced. The pressure for space in our journals has also decreased the number, as many editors believe that the length of articles is needlessly increased by the use of both. It has, therefore, become a matter of some concern to know whether it is better to report the results in tabular or graphic form.

There is no question as to which method is the more accurate, *i.e.* which permits the fuller presentation of the results. The data—their means, measures and ranges of variability, significance and reliability—are usually presented in greater detail in tables than in line-cuts. They may be presented in tables to as many decimal places as desired, but in graphs the unit-digit, not to mention decimal places, are usually obscure.

"Better" in this connection does not necessarily mean, however, more 'detailed' or 'accurate.' It may mean 'more clearly' or 'more convincingly.' Trends, relations, and functions are more clearly, forcefully, and convincingly shown in line-cuts than in tables. Only the rare reader, one who can graphically visualize the data given in tables, will be able clearly to detect trends and relationships without plotting the graphs.

Which method, then, is the better in a given case? Though the problem is obvious, a direct experimental attack upon it has not, as far as an extensive bibliographical search reveals, been made. This is doubtlessly due to the complexity of the problem—to the difficulty involved in designing a procedure that would be free from individual differences and prejudices.[1] What has not been accomplished by a direct experimental attack may, however, be accomplished by indirection; *i.e.* by observing what authors select from another study to illustrate their text- and handbooks. They surely must be credited with selecting what they regard as the better of the two methods of presenting the results.

Nearly forty years ago, an experiment on "Obliviscence during sleep and waking,"[2] was reported in which its results were presented in both tabular and graphic form. This study has been widely quoted and every author reporting it has requested permission to reproduce the line-cut—not one, to reproduce the table. This can only mean that, in this instance in which the authors have a choice, they regarded the graphic method of presenting the results as superior to the tabular.

TABLE I

NUMBER OF SYLLABLES CORRECTLY REPRODUCED AFTER VARIOUS INTERVALS (N=8)

		Interval							
O	Condition	1 hr		2 hr.		4 hr.		8 hr.	
		Av.	*mv.*	Av.	*mv.*	Av.	*mv.*	Av.	*mv.*
H	sleep	7.1	1.6	5.4	1.2	5.3	0.4	5.5	2.2
	waking	4.4	0.9	2.8	0.8	2.4*	0.8	0.4	0.4
Mc	sleep	7.0	0.7	5.4	0.6	5.8	1.2	5.8	1.2
	waking	4.8	1.2	3.4	1.4	2.1*	0.4	1.4	0.9

* Values that were incorrectly reversed in original graph.

The line-cut in this article contains a slight error—the points in the curves at the 4-hr. waking-period were reversed for the two *O*s. Figure 1 shows the graph published in the original paper; Table I, the data from which the graph was plotted;[3] and Figure 2, the graph as it should have been drawn. The error is obvious, but it was not detected until 1932—eight years after publication.[4] After the error was detected, it was, of course, brought to the attention of everyone requesting permission to reproduce the graph. Not one of those to whom the error was pointed out mentioned it in his request to reprint and not one stated, after it was pointed out, that he had previously detected it.

---

[1] A questionary, frequently resorted to in such instances, would merely summarize the prejudices at the moment of those answering it. What a correspondent thinks he prefers does not correlate highly with his preferences in a life-situation. Cf. E. G. Boring, When and how to reply to a questionary, this JOURNAL, 37, 1926, 612 f.

[2] J. G. Jenkins and K. M. Dallenbach, *op. cit.*, this JOURNAL, 35, 1924, 605-612.

[3] This table is a simplification of Table III in the original (*op. cit.*, 609). It was recast to bring out clearly the points here at issue.

[4] The error was discovered by a student of mine who reproduced, with corroborative results, the original study by the method of savings. He plotted the curves from

Why was the error so long in being detected? And why, since its detection, has it not been repeatedly detected? We shall never know the answer to these questions, but my guess, based upon the unanimity of authors' choice of the graph mentioned above, is that a graph gives at a glance so much information and attracts so much attention to itself that the table, from which it was plotted, is ignored or only casually inspected. Certainly few readers, only those with especial interest in a paper,[5] can afford the time required to study a table when a graph is available. All the evidence obtained from the reproduction of the study mentioned here indicates that the graphic method is 'better' than the tabular. Tables, since graphs are based upon them, are necessary, but they are like background rocks, heavy and uninteresting. Graphs, on the other hand, spice the reports; clarify them, and make them interesting and palatable.

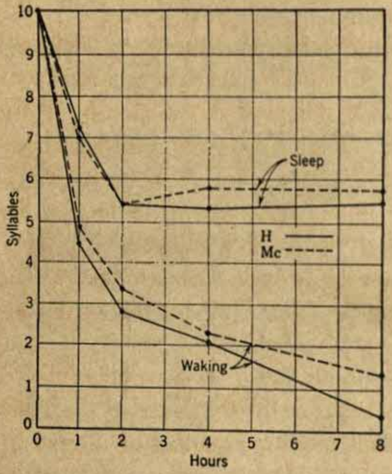


FIG. 1. THE GRAPH PUBLISHED

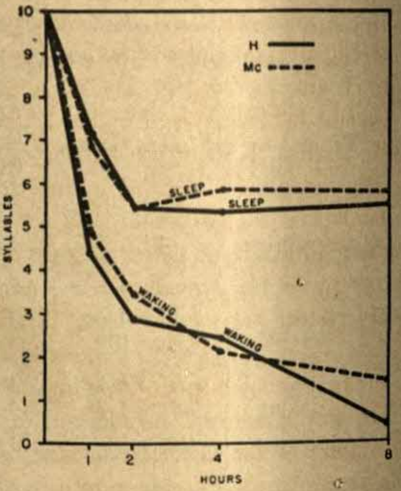


FIG. 2. THE GRAPH CORRECTLY DRAWN

The University of Texas KARL M. DALLENBACH

---

the table and was surprised to discover an error in the original graph at the 4-hr. waking-period—as was I when he brought it to my attention. Cf. E. B. Van Ormer, Retention after intervals of sleep and waking, *Arch. Psychol.*, 31, 1932 (No. 137), 1-49; Sleep and retention, *Psychol. Bull.*, 30, 1933, 415-439, esp. 424.

[5] A case in point is Van Ormer's discovery of the error in the graph mentioned here. See footnote 4, above.

THE NINTH ANNUAL MEETING OF THE SOUTHEASTERN PSYCHOLOGICAL ASSOCIATION

The ninth annual meeting of the Southeastern Psychological Association was held jointly with the Florida Psychological Association and the Florida Association of School Psychologists in Miami Beach, Florida, on April 25-27. At the joint meetings, 96 papers, 14 symposia, 3 invited addresses, and 2 presidential addresses were presented.

The presidential address of the Southeastern Association was delivered by Irwin A. Berg of Louisiana State University on "The generality of deviant behavior," and of the Florida Psychological Association by Byron B. Harless, Kirkpatrick Associates on "Nothing happens until somebody sells something." Of the invited addresses, two were given by Ralph H. Tindall of the Milwaukee, Wisconsin, Public Schools, on "A look at the expanding role of the school psychologist" and "Pros and cons of psychotherapy in the school setting," and one by Morton Deutsch of the Bell Telephone Laboratories on "Psychological aspects of international conflict."

The annual business meeting of the Southeastern Association followed Dr. Berg's address. In addition to the selection of the place of the 1964 meeting by the Executive Committee and the announcement of the newly-elected officers, the membership voted: (1) to limit meetings to places that are desegregated; (2) to express disapproval of the policy of the American Psychological Association of charging regional associations for the publication of their Proceedings; (3) to urge its members to oppose restrictive legislation relating to animal experimentation; and (4) to honor Lula Stevens, Mississippi College for Women, retired, by election to Life Membership.

The officers elected for the year 1963-64 are as follows: President, Susan W. Gray, George Peabody College. President-Elect, Louis D. Cohen, University of Florida; Past President, Irwin A. Berg, Louisiana State University. Ralph Mason Dreger, Jacksonville University, continues his service as Secretary-Treasurer for another year. Wilse B. Webb, University of Florida, was elected to a three-year term as a member-at-large of the Executive Committee. The other members of the Executive Committee are: Julius Seeman, George Peabody College, and Theodore Landsman, University of Florida. The 1964 convention is to be held at Gatlinburg, Tennessee, April 2-4. The rolls of the Association now number 629 Members and 95 Student Affiliates.

Jacksonville University, Florida RALPH MASON DREGER

FIFTY-NINTH ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The fifty-ninth annual meeting of the Society of Experimental Psychologists was held at Brown University on April 1 and 2, 1963. Harold Schlosberg, Chairman of the Society for the year, was the Chairman of the meeting.

Members and fellows present were: Chapanis, Estes, Fitts, Garner, Geldard, E. Gibson, J. Gibson, Grant, Helson, Humphreys, Hurvich, Irwin, Kimble, Kendler, Mackworth, Marquis, Meyers, Miles, Mueller, Neff, Newman, Pratt, Ratliff, Schlosberg Solomon, Spence, Stellar, Stevens, Riggs, Rosenblith, Underwood, Volkmann, Wever and Wickens. Gunnar Johansson was an invited guest.

P. Teitelbaum, W. J. McGill, R. D. Luce, and C. P. Duncan were elected to membership in the Society, bringing the total membership to 66 members and 21 fellows.

Members presented reports of research in progress at scientific sessions held the afternoon of April 1 and on April 2.

The Warren Medal for 1963 was awarded to W. K. Estes "for development of a mathematical theory of learning which he has painstakingly tested and which has proved to be a powerful tool for analyzing the role of stimulation in both human and animal learning."

The Society accepted the invitation of Stanford University and the University of California, Berkeley, to meet at these universities in 1964. F. A. Beach was elected Chairman for 1963-64.

Princeton University JOHN L. KENNEDY

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. By FRANK ROSENBLATT. Washington, D.C., Spartan Books, 1962. Pp. xvi, 616. $6.50.

In the words of the author, ". . . our objective has been to discover a physical system, or abstract model, which will be capable of "perceiving" its environment, and learning to recognize those objects or events which it has perceived in the past. However, since it is our purpose to understand the actual mechanisms employed by the brain, rather than simply to construct a new type of computing device, the perceptron models are constrained in their organization and dynamic properties by what is known of the biological nervous system. Rather than attempting to "invent" or "construct" a machine which will calculate such things as similarities or geometrical properties of stimuli, the approach has been to begin with an hypothetical network of idealized neurons, or nerve cells, resembling the brain in its general organization, and then analyze the system mathematically to determine whether or not it possesses "psychological" properties of interest. Where the model is found to deviate markedly from the behavior of biological systems, modifications are suggested, and the new model that results is subjected to the same sort of analysis. In this fashion, it is hoped that the necessary conditions for a system to "perceive" in the same manner as the brain can be abstracted" (p. 574).

Among the neural net models recently reviewed by Uhr in the *Psychological Bulletin,* Rosenblatt's perceptrons are distinguished by their emphasis on randomness in the arrangement of elements comprising the hypothetical nerve nets. In this approach, Rosenblatt is exploiting ideas widely discussed in psychological circles upon the publication of Hebb's *Organization of Behavior* in 1949 and an approach to automatic pattern-recognition first attempted by Farley and Clark in 1954. The advent of the high speed general purpose digital computer has made it possible to simulate such networks and to ascertain some of their characteristics by varying their structure and exposing them to various environments. The results of several such simulations are summarized in this volume so that it is fair to say that the perceptron models at this writing have received more systematic attention than any of the other nerve net models which have been proposed to deal with discriminative and recognitive behavior. Many who were baffled or enraged at some of the early expositions of the perceptron either in print or from the podium will find that the claims for it in this volume are considerably less inflated and the description, while something less than lucid, is nevertheless sufficient to enable one to write programs to simulate given perceptrons closely enough to replicate results reported in the book.

Though Rosenblatt goes to considerable pains to insist that his perceptron models "are constrained in their organization and dynamic properties by what is known of the nervous system," this constraint is not great. In none of the models discussed in detail are the temporal characteristics of neurone firing embodied, and in picking parameters to determine the statistical characteristics of the various networks discussed, it is rarely neuroanatomy or neurophysiology which supplies the threshold values of an A unit, or the proportion of inhibitory and excitatory connections between A units and S units. Like others who have been concerned with neural net models, Rosenblatt uses only the logical properties of the neurone in defining the elements which comprise his networks. Such simplification may be necessary in order to construct a model which can be simulated on existing computers, but until such programs embody considerably more of what is known of the characteristics of individual neurones, their arrangement and distribution, they will be of little help to the neurophysiologist in testing hypotheses about the functioning of a given organism's nervous system, nor will the results of monte carlo runs with such models provide the neuroanatomist with a compelling reason to develop new data in his field.

Whether or not one regards the perceptron as a valid neurophysiological model, one may ask the extent to which it provides a useful means for understanding the discriminative behavior of organisms. The answer to this question is provided in large part by the author's remark on page 71 that "Perceptrons considered to date show little resemblance to human subjects in their figure-detection capabilities, and Gestalt-organizing tendencies." This failure stems almost entirely from the perceptron's dismissal of all topological information about the pattern illuminating the retina. The random connections between the retina and the association units preclude any feature extraction which depends upon the geometry of the pattern. This being so, it is at first surprising to find the perceptron at all capable of discriminating two forms which differ only in their geometry. In Experiment 7 Rosenblatt shows that a simple perceptron can indeed discriminate between a square and a triangle, each of which illuminates 20% of a 20 × 20 retina. After 1000 trials, this perceptron correctly discriminates these two figures with an accuracy of 74%. A persistent problem with such results is that no indication of the variability in the system, nor of its asymptotic performance level is given. In the case of the simple perceptron employed in Experiment 7, these are easily determined. One such perceptron simulated on an IBM 7090 yielded proportions of correct discriminations for a square-triangle problem which fell between 62% and 82% on the tenth block of 100 trials when 15 different runs were made on the same perceptron. The mean was 70%—very close to the mean reported by Rosenblatt. With the same figures, but using 15 different perceptrons rather than 15 runs on the same perceptron, the percentages ranged from 50% to 88%. On this particular kind of perceptron, and using stimulus-controlled reinforcement, the mean asymptotic performance level may be determined simply by presenting all possible figures in all translations in a training series, and then seeing how many of these are subsequently identified correctly. On the 15 perceptrons mentioned above, these mean asymptotes ranged from 76% to 90% with a mean of 80%. None of the 15 perceptrons had reached its asymptote by the tenth block of 100 trials.

Though the perceptron is singularly insensitive to the arrangement of points comprising a pattern, it is very sensitive indeed to size and to the relative number of times each figure is presented for discrimination. On page 171, Rosenblatt shows the results of an experiment in which one pattern was shown four times as often as the other, with the result that after 200 trials the perceptron was no longer making any discrimination at all, but simply identifying everything presented to it as the more frequently presented figure. In the simple perceptrons, and even in some of the more complex ones, relative size of the two figures to be discriminated plays the same role, the larger figure gradually extinguishing any response to the smaller figure.

Though the author presents certain mathematical properties of his perceptrons in some detail, he gives little information about their expected behavior in various environments. For example, most of the experiments reported deal with figures which illuminate a considerable fraction of the retina, and it would be of some interest to know how well the perceptron could do with smaller figures. In ordinary reading of the printed page, we recognize letters whose images illuminate approximately 1% of the fovea. The perceptron described in Rosenblatt's experiment 7 was presented with a discrimination of a square and triangle which illuminated 6% of the S units. With such a small figure asymptotic performance level for this perceptron was a mere 59% correct.

With these limitations, the psychologist may well wonder if the perceptron offers a model of human pattern recognition and discrimination which is worth further exploration. Rosenblatt believes that by placing constraints on simple perceptrons whose capabilities are explored in this volume, most of their limitations may be overcome; and he even admits (p. 499) that "a passive filter-type system (such as a simple perceptron) cannot be designed which will economically recognize topological abstractions and relations. . . ." However, in this volume such more complicated perceptrons are merely suggested, and experimental results indicating their actual capabilities are not given.

There is a growing awareness that the computer program offers a powerful language for the construction of models of many human psychological processes: recognition, problem solving, rote memory, etc., and in fact this kind of endeavor already is called "artificial intelligence." Such programs offer a tremendous opportunity to psychology, but a word of caution is perhaps in order. The objective of the whole endeavor is easily lost when the program becomes an end in itself. To some extent this appears to have happened with the perceptron. As the perceptron program now stands, it is an unconvincing neurological model since few of its parameters are firmly rooted in neuroanatomical or neurophysiological data, and it offers little to psychologists who expect a model tailored to the known facts of human pattern recognition and discrimination.

A word should be said about the format of this volume. Produced by photo offset, it bears the marks of a hasty conversion of an earlier technical report. One is confronted immediately with an errata page whose last entry refers to a page in the book which doesn't exist.

Lincoln Laboratory
Massachusetts Institute of Technology

B. W. White

Introduction to Psychological Research. By William A. Scott and Michael Wertheimer. New York, John Wiley and Sons, 1962. Pp. xiii, 445. $7.95.

This book, the authors remark in their preface, grew out of a course for beginning graduate students at the University of Colorado. The purpose of the course and presumably of the book is, "to acquaint beginners with the practicalities of doing original research."

The book seems to have fulfilled its purpose well. It is, in the reviewer's opinion, the most useful book on psychological research, at its level, available today. The authors carry the student through the entire research process. Representative chapter headings are the following: initial stages on a research project, conditions for generalizing findings, control over research variables, developing quantitative measures, instrumentation (verbal measures, equipment, human judgment), sampling subjects, experimental design, dealing with subjects, data collection, data analysis, and interpreting and reporting the results.

The book reflects several unusual topics. Many chapters have supplements, usually on statistical matters, but there are also supplements on the major psychological journals (chapter 2), on sources of financial support for psychology (chapter 1), and on types of psychological equipment (chapter 7). The chapter on equipment is noteworthy in that it spells out certain problems in experimenting with equipment such as the "warm-up effect," "instrument drift" and constant errors, applying to equipment the same standards that apply to psychological tests or other observations of behavior. The student is even warned about "falling in love with equipment."

Written as the book is by a social psychologist (Scott) and an experimental psychologist (Wertheimer), there is a broad "social science" perspective throughout. In particular, the problems of sampling of subjects (sampling design) is well placed in a book for general use by graduate students; these problems are usually more discussed by social psychologists and sociologists than by experimental or clinical psychologists. The problems of working with subjects are ably discussed in chapter 11, "obtaining and dealing with subjects." This chapter is the reviewer's favorite and should, in his opinion, be required reading of all graduate students, particularly the section on maintaining cooperation with subjects and the one on ethical considerations in using subjects.

The book does have three shortcomings, shortcomings with which others may disagree or even find to be advantages. The first of these concerns the introduction of certain statistical procedures. The authors introduce many statistical tests into the text, and, particularly, into the supplements at the end of chapters. While some of these are usual (Chi-square), most are somewhat special (measuring uncertainty reduction). Whatever the merits of these statistical procedures, and they range from illustrations or applications that most graduate students can understand to techniques that would require much class time to explain, the result is a hodgepodge intermingling of statistical procedures and research methodology. Of course, one can argue that research cannot be understood without statistics, but the impression is one of bringing in "pet" ideas rather than that of a serious attempt to think through and logically relate research methodology and statistics.

A second shortcoming revolves around their discussion of the usual scales of measurement (nominal, ordinal, interval, ratio). They admit that most psychological data are, at best, ordinal. In discussing correlation, they do not recommend product-moment correlations unless one has interval data; otherwise for ordinal data, one should use τ or h. But everyone knows that factor analysis has correlated all sorts of data which, whatever its merit, is highly repeatable. The relation of the rules of statistics to actual problems is undoubtedly not understood by the statisticians themselves, nor can it be. But the barrage and counterbarrage in psychological journals of appropriate tests and appropriate assumptions have serious consequences in the laboratory. The researcher finds himself damned if he uses a simple statistic like Chi-square when a more complicated one could be used, and doubly damned if he uses a complex statistic when he may have violated an "assumption." The end result is that graduate students (and others) spend literally weeks running statistical tests, agonizing over what to use and why, when they could have spent the time doing a meaningful experiment. (No one wants to be cited as a "horrible example".) All of this is not Scott and Wertheimer's fault, and they are very aware of the problem (see vii and p. 122), but they do, to some extent, contribute to the chaos by citing the ordinal-interval-ratio scale distinctions without discussing in detail the consequences of violation of the various "assumptions" on conclusions reached through the statistical techniques. The reviewer feels that their treatment, despite occasional "don't worry" comments, will leave the student confused and with no firm guide lines to get on with the research.

Third, the authors are well aware of the fact that research often does not follow the "rules." They quote in the introduction a passage from Hebb on the dangers of research methodology in a vacuum and emphasize there that the research process in action is not standardized. They do not, however, emphasize this enough throughout the book. The authors warn about "falling in love with equipment," as has been mentioned, but they also might have warned about the dangers of design-ascetic research that is devoid of thought ("shot-gun empiricism—factorial style"). As psychologists, perhaps in our search for "science," we have often yielded to *the illusion of precision* produced by electronic equipment, mathematical equations, or complex designs. These are matters that books on research methodology must face; they must balance presentations of method with more warnings about method for method's sake, and, in so doing, show that the *problem* is more important than the *methodology*.

In sum, this is an excellent book to introduce the methods of psychological research to beginning graduate students. Presumably the book will be supplemented, in most graduate programs, with an introduction to actual research. The shortcomings listed are no barrier to using the book as a text. Indeed, one semester's use here has been well received by students, a better reception than two previous texts. Students need to know research methodology (a firm grounding in research methodology is not a handicap to original research, in the reviewer's opinion) and this book is a fine introduction to it.

RICHARD D. WALK

George Washington University

Psychological Differentiation: Studies of Development. By H. A. WITKIN, R. B. DYK, H. F. FATERSON, D. R. GOODENOUGH, and S. A. KARP. New York, John Wiley and Sons, Inc., 1962. Pp. v, 418. $7.95.

This important book is "must" reading for all who are interested in individual differences, child development, and personality organization. The authors describe a large group of studies based upon and greatly expanding the earlier work on "field-dependence and field-independence" described by Witkin and associates in the 1954 book, *Personality Through Perception.* As implied by the title, this dimension of individual differences has been reconceptualized in terms of "psychological differentiation" on the basis of results of a veritable host of studies by the Witkin group, by graduate students, and by other investigators.

The extensive empirical work described provides impressive evidence that one form of articulation of experience increases progressively during development and that enduring individual differences are demonstrable in children's performances in laboratory perceptual tests, the Rorschach test, the TAT, self-drawings, conformity experiments, certain subtests of intelligence tests, measures of the sense of separate identity, active attitude, etc. Mothers' characteristics, test performances, and patterns of interaction with their children are shown to be related to their children's performance in criterion tests of articulation of experience. Case studies of clusters of extreme subjects round out the evidence indicating that this dimension of personal organization cuts across the traditional categories of intelligence, perception, memory, etc.

The conceptual scheme offered may be somewhat less impressive than the remarkable evidence of individual consistencies presented. The term "psychological differentiation" seems to imply more generality than is warranted even by the notable consistencies described. The most generally used criterion procedure—Witkin's Embedded Figures Test—is, as the authors note, demonstrably interchangeable with the embedded figures tests Thurstone used to measure "flexibility of closure," an intellectual ability long known to be but one of a considerable group of capacity or ability variables. Chief among these is a highly generalized verbal ability, which, as the authors note, is not related to articulation of experience as measured by them.

The overgeneralization involved in referring to the articulation of certain aspects of experience as psychological differentiation may also have led to the authors' impression that such articulation is, or should be, linked to intellectualizing. Intellectualizing, however, is typically rooted in the very kinds of verbal skills that are uncorrelated with this kind of articulation. The articulation of experience described is, in fact, correlated most highly with the Picture Completion, Block Design, and Object Assembly subtests of the WISC, which have repeatedly loaded on "performance" rather than verbal factors.

The authors refer to studies by Steven G. Vandenberg (further substantiated by his recent work) suggesting that a significant genetic factor contributes to the individual differences described in this book. The genetic hypothesis seems not to have been given adequate representation, however, in the design of the studies reported on the origins of this type of articulation. To the writer, the unusually high reliabilities of criterion scores over time and the intellectual-ability characteristics of the form of articulation explored also suggest that a genetic factor may be a major determinant of these individual differences. Some of the results that appear to be caused by features of the mother-child interaction may, in fact, be primarily attributable to such a genetic factor.

The most questionable formulations in this generally excellent book concern relations between the kind of articulation studied and defense mechanisms. After collapsing defense mechanisms into groups they describe as more primitive versus more refined and specialized—a procedure of dubious validity—the authors indicate that ratings based on this scheme support their hypothesis concerning the relationship of articulation of experience to defensive structure. Amalgamating different defenses in this way has the unfortunate effect, however, of making the "defense" ratings more similar than ratings of particular defenses to measures of articulation *per se*. The ensuing inference that children high in articulation tend to be obsessive-compulsive, whereas children low in articulation tend to be "hysterical," seems unjustified on the basis of these results. In reporting evidence in support of the latter hypothesis, the authors ignore published evidence suggesting that when defenses are rated separately, the defense of isolation (presumably the key to obsessive-compulsive behavior) is unrelated to Witkin's criterion measures of articulation. The authors' discussion of projection as a highly differentiated, structured defense is also surprising. A large body of psychoanalytic work delineates projection as one of the two or three most primitive defenses (although its superficial expressions may appear complex in particular cases). Developmentally and structurally, projection is probably less complex and differentiated than repression, which is described by the authors as primitive. One must distinguish between a defense mechanism and its consequences. The study of the relationship between articulation and the single defense of denial as manifested in the TAT, however, is understandable and useful.

As noted above, psychological differentiation may have a number of different forms. It is worth noting, in this connection, that published studies of the kind of articulation described by Witkin *et al.* (referred to in terms of this or other rubrics) further substantiate the lack of universality of this dimension. Baggaley has shown, for example, that this kind of articulation is relevant to concept formation only under special conditions. The writer and his associates have repeatedly shown that under other conditions articulation as measured by the Witkin procedures is unrelated to spontaneous conceptual articulation. In the case of learning and recall, Long recently showed that articulation as described by Witkin *et al.* allows prediction of performance in a special kind of verbal learning test. When the most common forms of verbal learning tests have been used, however, results have been ambiguous or negative, as the authors point out.

The critique of the authors' conceptualization offered here is in no sense intended to detract from their notable achievements in demonstrating the generality of individual differences in one dimension of personal organization. The quality and quantity of the empirical work they summarize are truly commendable. The reader will find himself eager for reports of further studies by this research group.

RILEY W. GARDNER

The Menninger Foundation

Television in the Lives of Our Children. By WILBUR SCHRAMM, JACK LYLE, and EDWIN B. PARKER. Stanford, Stanford University Press, 1961. Pp. vii, 324. $6.00.

The study *Television in the Lives of Our Children* can be reviewed from many points of view. The report of three years of research on 6000 children and 2300 parents, teachers, and school officials presents many detailed descriptions and tables of the viewing habits of children. The reviewer has chosen to suggest the extent to which the material tends to reinforce some very general concepts of growth and development rather than to recite detailed findings. The analysis is, of necessity, sketchy.

The concept of multiple causation. Many of the mistakes in the interpretation of the first studies of the impact of television were produced because of a lack of control of the factors in the samples of the population studied. The minimum controls supported by the present study would examine variables such as intelligence, social background, age, and sex. The amount of viewing appears to reach a peak between ages 11 and 13 and then to decline. Sex differences are more qualitative than quantitative. Viewing will vary according to level of ability. Patterns of light or heavy viewing tend to go through a family. The complexities of the emotional and social aspects require a clinical analysis by a human correlator.

The concept of individual differences. Quantitative and qualitative differences are amply documented. Thus in a particular grade some children will average more than four hours per day of viewing and some less than fifteen minutes per day. Similarly, but in a more subtle and complex fashion, the effects of television are obscured by the characteristics within the individual. There is a suggestion to the reviewer in the material that one might construct a "viewing age." Anyone interested?

Seeking and self-selection. There is evidence that children seek according to maturity and needs from the environment supplied by the television screen. Thus able children seek the experience earlier, reach a peak earlier, and go into an adult phase earlier. Some need refuge and escape in fantasy and others are more oriented toward reality. Immediate and deferred satisfactions follow the known linkages with ability and social class.

Presence and absence of experience. It is an old generalization that we "learn our experiences." It is logical that television viewers learn some things not learned by non-television viewers. Thus contrast in viewing in a town with television and a town without television indicates early but transient superiority in vocabulary, and subsequent equivalence as experience is supplied through other media. This tendency coincides with most studies of early stimulation. The specific qualitative details of information obtained are related to their emphases in the media to which the children are exposed.

Homeostasis. Some balance of activities exists within the total time and energy of the child. Thus the same ends may be obtained with different means in a town without television by more extensive use of comics, movies, and radio. Availability of television reduces play time somewhat, homework slightly, and delays bed-time by thirteen minutes.

Frustration, regression, and aggression. Interiorization and externalization of tensions appear to occur in complex and individual ways in accord with classic theory. Uniform affective and behavioral effects as in delinquency and disturbance are not supported. The report concludes with psychiatric comments giving more questions than answers but illustrating the complexities of the emotional life.

Summary. Anyone who is inclined to accept the generalization of a "partialist" on the effects of television or who wishes to be informed in depth should read the report. The perspective of the study places claims and counter-claims in proper balance without minimizing the pervasive and influential role of television in our culture.

University of Michigan WILLARD C. OLSON

Explaining "Teaching Machines" and Programming. By DAVID CRAM. San Francisco, Fearon Publishers, 1961. Pp. vii, 86. $2.00 (paperback)

Any book describing an instructional approach as new and novel as programming not only should present the essential facts and background information, but it is almost obligated to utilize programming principles in assisting the reader's learning as well. In his brief volume, Cram provides an introduction to programmed instruction designed to inform lay readers about the characteristics of programmed materials, with special emphasis on the distinction between the linear programs advocated by B. F. Skinner and the branching programs developed by Norman Crowder. While the content of the book is elementary, perhaps more so than necessary, it does touch upon many facets of interest to a potential program user. The book's outstanding feature, however, is that it consists of a series of programs which include both linear and branching segments. In presenting information this way, Cram avoids the frequent criticism of using a conventional textbook format to explain the instructional superiority of programmed materials.

Unlike a text, a program can be examined with respect to teaching effectiveness as well as content. In this sense, the information contained in a program is of no particular importance unless students can be expected to learn from it. Regardless of whether a linear or branching format is used, programs can be tried out and revised as often as necessary to insure learning will occur and that the desired objectives will be attained. Since such procedures are an essential step in the writing process, a program can be assessed more directly than most other instructional materials by establishing its effectiveness empirically. Cram has made such an evaluation of his book possible by providing a self-test at the end of each chapter. The resulting opportunity to assess a program as one should be assessed could not be passed up.

With the assistance of Alex W. Andres, the introduction and first chapter of the book, consisting of 13 pages in scrambled-text or branching format, were administered to 31 undergraduates enrolled in an educational psychology course at the University of Pittsburgh. The "self-test" covering the salient points in this section of the program consisted of four questions. The first three of these required essay responses while the fourth involved a short answer. Although no explicit key was given, liberal grading standards were prepared on the basis of the information contained in the section which gave at least partial credit for even an approximation of the correct answer. Since the questions were not of equal complexity, weights were given by assigning points to partial answers. The highest possible score was eight.

The results indicated that only two of the 31 students obtained a perfect score. One student had zero correct. The mean score was 5.1, or 64 percent. No part of any question was correctly answered by all students, a finding not particularly attributable to scoring rigor since the most objective item was satisfactorily answered by the fewest students, ten. Because a branching technique was employed, the scores of the 18 students who used at least one branch were compared with the scores of the 13 students who did not avail themselves of the additional material. The students who branched averaged 4.5 while the students who did not averaged 6.0.

These results speak for themselves. Comparing them with the increasingly popular interim standard for programs of 90 percent or better achievement by 90 percent of the students suggests that at least this section of the Cram program is not particularly illustrative of the high quality of instruction attainable with programmed materials. While one can only guess at what the outcome would be had a more conventional text been assessed in this manner, it is regrettable that this book, despite its format, fails to accomplish its goal of producing a more thorough understanding of programming than it apparently does.

American Institute for Research DAVID J. KLAUS

The Psychology of Aggression. By ARNOLD H. BUSS. New York, John Wiley and Sons, 1961. Pp. x, 307. $7.95.

Arnold H. Buss of the University of Pittsburgh has presented his research colleagues with a useful review of psychological studies of aggression in humans.

One's initial doubt that enough has been done to fill a book is quickly dissipated; although far from exhaustive in its coverage, the 300-page treatment is crammed with synopses, reviews, and summaries. As a compendium, Buss's book has been welcomed by research workers and their graduate student apprentices, who find in it extended discussion of both laboratory studies of situational determinants of aggression and psychometric studies of aggression as an enduring response tendency. Also included in the coverage are studies of selected psychosomatic disorders, of aggression in children, and of prejudice as a form of socially institutionalized aggression.

Not surprisingly, the book does not leave its reader with a coherent view of aggression in its diverse manifestations. There are at least three reasons for this, none of which reflects adversely on Buss's adequacy in accomplishing the task he set for himself.

One is the methodological unevenness of the studies which are reviewed. Given the state of the field, the standards of methodological adequacy employed by the author in selecting studies for review could not be stringent, and included in the book are many studies whose design is so insufficient or whose measures are so gross that little confidence can be placed in the findings.

Second, the studies were pursued in a variety of different theoretical contexts. Drawing as he does on studies embedded in the biological tradition of medical research, in the S-R allegiance of the psychological laboratory, in the cognitive theories predominant in social psychology, and in the psychoanalytic view held in the clinic, the author is hardly able at once to be faithful to his sources and at the same time to provide a satisfyingly unified account of aggression.

Finally, on the face of the matter it is not evident that aggression is a sensible topic for separate treatment. In psychology as in other fields, a perennial problem is how to slice the pie. The two aims of the slicer—to withdraw a unified piece of edible size and to avoid doing violence to the remaining confection—must also be the goals of anyone attempting to partition the burgeoning, sprawling field of psychology into manageable units. One may doubt that aggression as a topic has either the coherence or the extent to justify its separation, and one may also wonder whether justice can be done to the psychology of aggression when it is dealt with as a topic abstracted from the corpus of psychology. The experience of medical research has been instructive. Explication of a particular symptom (headache, fever, etc.) in all its varieties and subtleties of instigation and manifestation has not often led to fundamental understanding of any of the several distinct and essentially unrelated illnesses from which this single symptom may arise. Similarly, it is not clear that focusing on aggression per se—its development, forms, correlates, and instigators—is the optimum strategy for arriving at an understanding of such fundamental topics as say the nature of motivational systems, the biological bases of social behavior, psychosomatic interactions, and the like.

Center for Advanced Study
in the Behavioral Sciences

ALBERTA ENGVALL SIEGEL

Tests in Print. By O. K. BUROS. Highland Park, New Jersey, The Gryphon Press, 1961. Pp. xxix + 479. $7.00.

What can the poor reviewer say about a book of this type? He cannot say it is interesting reading, because it is not intended to be. He cannot say that it is accurate, because the immense number of references defies anything other than a superficial spot checking. With an encyclopedic book of this type, perhaps all the reviewer can, or should do, is to discuss three points: (a) what the book contains, (b) what purpose the book is intended to serve, and (c) how well it seems to serve that purpose.

The book is essentially a list of every test published in English which the author could find in an exhausive search of the *Mental Measurement Yearbooks,* test catalogues, professional journals, and other sources. The list is divided into two parts, consisting, respectively, of "tests in print" (2126 entries) and "tests out of print" (841 entries). For each entry the major pieces of information supplied are: (a) test-authors, (b) publisher(s), (c) cross-references to reviews in *MMY* (when available) and other places, (d) names of parts and subtests, (e) dates of publication, (f) age-levels appropriate for each test, and (g) "special comments," *e.g.* "new form issued annually." The following excerpt illustrates the listings (pp. 301):

> *Study of Values: A Scale for Measuring the Dominant Interests in Personality, Third Edition.* Grades 13 and over; 1931-60; 6 scores: theoretical, economic, aesthetic, social, political, religious; Gordon W. Allport, Philip E. Vernon, and Gardner Lindzey; Houghton Mifflin Co.
>
> For additional information and a review by N. L. Gage of the Second Edition, see 5:114 (57 references); for reviews by Harrison G. Gough and William Stephenson, see 4:92 (25 references, 1 excerpt); for a review by Paul E. Meehl of the original edition, see 3:99 (61 references).

An introduction provides (a) an overview of the purpose and scope of the book, (b) statistical breakdowns of the types of tests included, and (c) some sardonic, bittersweet comments about the dogmatically superstitious behavior of the average test user. Following the lists of tests are verbatim reprints of *Technical Recommendations for Psychological Tests and Diagnostic Techniques and Technical Recommendations for Achievement Tests.*

How will the book be useful? It will be useful for locating critical reviews for particular tests. The book would be valuable if for no other reason than serving as a master index to reviews in the five *Mental Measurement Yearbooks.* It will be useful for the person who, with no particular test in mind, is searching for those available in a particular area, *e.g.* measures of socioeconomic status. Finally, the book will be valued by those omnivorous fact-gatherers who would, for example, enjoy remembering that a Hebrew Aptitude Test is available.

Who will want the book? It will be a valuable reference book for university libraries. In addition, the book inevitably will rest on the bookshelves of all those who deal extensively with tests: measurements specialists in research centers, school programs, clinics, industry, and the armed forces. Now that the book is available it is hard to see how we ever did without it.

Vanderbilt University JUM C. NUNNALLY

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

BAKER, ROBERT A. (ed.), *Psychology in the Wry.* Princeton, N.J., D. Van Nostrand Co., Inc., 1963. Pp. xii, 170. $1.75. (paperback)

BARKER, R. G., WRIGHT, H. F., BARKER, L. S., and SCHOGGEN, M., *Specimen Records of American and English Children.* Lawrence, Kan., Univ. of Kansas Publications, 1961: Pp. 271.

BARTON, R. F., *Autobiographies of Three Pagans in the Philippines.* New Hyde Park, N.Y., Univ. Books, Inc., 1963. Pp. xxxv, 271. $7.50.

BASSETT, WILLIAM T., *Counseling the Childless Couple* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1963. Pp. 139. $2.95.

BORING, EDWIN G., *History, Psychology, and Science: Selected Papers.* New York, John Wiley & Sons, Inc., 1963. Pp. xii, 372. $8.95.

BRAUN, JOHN R., *Contemporary Research in Learning: Selected Readings.* Princeton, N.J., D. Van Nostrand Co., Inc., 1963. Pp. vi, 229. $1.95. (Paperback)

BROWN, ROBERT, *Explanation in Social Science.* Chicago, Ill., Aldine Publ. Co., 1963. Pp. viii, 198. $5.00.

BUCKNER, DONALD N., and NCGRATH, JAMES J. (eds.), *Vigilance: A Symposium.* New York, McGraw-Hill, 1963. Pp. viii, 269. $8.95.

CAPRIO, FRANK S., and BERGER, JOSEPH R., *Helping Yourself with Self-Hypnosis. A Modern Guide to Self-Improvement and Successful Living.* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1963. Pp. xiv, 210. $4.95.

COOPER, JOSEPH B., and MCGAUGH, JAMES L., *Integrating Principles of Social Psychology.* Cambridge, Mass., Schenkman Publ. Co., Inc., 1963. Pp. xi, 320. $7.50.

CRUICKSHANK, WILLIAM M. (ed.), *Psychology of Exceptional Children and Youth.* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1963. Pp. xii, 623. $7.95.

DUNLANY *et al.* (eds.), *Contributions to Modern Psychology* (2nd ed.). New York, Oxford Univ. Press, 1963. Pp. xii, 484. $3.50. (paperback)

FAUVILLE, ARTHUR, *Perception Tachistoscopique et Communication.* Louvain, Publications Universitaires, 1963. Pp. 56. 80 FB. (paperback)

FRANK, JEROME D., *Persuasion and Healing: A Comparative Study of Psychotherapy.* New York, Schocken Books, Inc., 1963 (Copyright 1961). Pp. xiv, 282. $1.95. (paperback)

FRASER, J. MUNRO, *Industrial Psychology.* New York, Macmillan Co., 1963. Pp. ix, 181. $2.45. (paperback)

FREUD, SIGMUND, *Collected Papers* (First Paperback Edition Complete in Ten Volumes). New York, Collier Books, 1963: *Dora: An Analysis of Hysteria.* Pp. 157. $.95; *Early Psychoanalytic Writings.* Pp. 254. $1.50; *General Psychological Theory.* Pp. 224. $.95; *The History of the Psychoanalytic Movement.* Pp. 125. $.95; *Three Case Histories: The "Wolf Man," the "Rat Man," and the Psychotic Doctor Schreber.* Pp. 319. $1.50; *The Sexual Enlightenment of Children.* Pp. 189. $1.50; *Sexuality and the Psychology of Love.* Pp. 223. $1.50; *Studies in Parapsychology.* Pp. 125. $.95; *Therapy and Technique.* Pp. 286. $1.50; *Character and Culture.* Pp. 320. $1.50.

FROMM, ERICH, *The Dogma of Christ and Other Essays on Religion, Psychology and Culture.* New York, Holt, Rinehart & Winston, Inc., 1963. Pp. x, 212, $3.95.

GOLDSTEIN, ARNOLD F., *Therapist-Patient Expectancies in Psychotherapy.* New York, Macmillan Co., 1963. Pp. xvi, 141. $4.50.

GREEN, BERT F., JR., *Digital Computers in Research.* New York, McGraw-Hill Book Co., Inc., 1963. Pp. xii, 333. $10.75.

GRINDER, ROBERT E. (ed.), *Studies in Adolescence.* New York, Macmillan Co., 1963. Pp. xii, 524. $4.50.

Grinker, Roy R., and Spiegel, John P., *Men Under Stress.* New York, McGraw-Hill Book Co., Inc., 1963 (Copyright 1945). Pp. xii, 484. $2.65. (paperback)

Grose, Robert F., and Birney, Robert C. (eds.), *Transfer of Learning.* Princeton, N.J., D. VanNostrand Co., Inc., 1963. Pp. vi, 194. $1.75. (paperback)

Hall, Calvin, *Dreams of American College Students.* Lawrence, Kan., Univ. of Kansas Publ., 1963. Pp. 217.

Hanfmann, Eugenia, Jones, Richard M., Baker, Elliot, and Kovar, Leo, *Psychological Counseling in a Small College.* Cambridge, Mass., Schenkman Publ. Co., Inc., 1963. Pp. xi, 313. $3.95.

Heckel, Robert V., and Jordan, Rose M., *Psychology: The Nurse and the Patient.* St. Louis, Mo., C. V. Mosby Co., 1963. Pp. 305. $5.85.

Johnson, James A., *Group Therapy: A Practical Approach.* New York. McGraw-Hill Book Co., 1963. Pp. xiii, 467. $10.95.

Judd, Deane B., and Wyszecki, Gunter, *Color in Business, Science, and Industry.* (2nd ed.). New York, John Wiley & Sons, 1963. Pp. x, 500. $15.00.

Kantor, J. R., *The Scientific Evolution of Psychology* (Vol.1). Chicago, The Principia Press, Inc., 1963. Pp. xxii, 387. $10.00.

Kirk, Samuel A., and Weiner, Bluma B. (eds.), *Behavioral Research on Exceptional Children.* Washington, D.C., The Council for Exceptional Children, 1963. Pp. xi, 369. $4.50, paper; $6.50 cloth.

Köhler, Wolfgang, *Intelligenzprüfungen an Menschenaffen.* Berlin, Springer-Verlag, 1963. Pp. 234. (Reissue of second edition, 1921). DM. 28.

Kuhlen, Raymond G., and Thompson, George G. (eds.), *Psychological Studies of Human Development.* New York, Appleton-Century-Crofts, 1963. Pp. xii, 638. $4.75. (paperback)

Latane, Henry A., Mechanic, David, Strauss, G., and Strother, George B., (Edited by Henry J. Leavitt) *The Social Science of Organizations: Four Perspectives.* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1963. Pp. ix, 182. $2.50.

Lewis, Donald J., *The Scientific Principles of Psychology.* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1963. Pp. xi, 580. $7.95.

Litterer, Joseph A. (ed.), *Organizations: Structure and Behavior.* New York, John Wiley & Sons, 1963. Pp. viii, 418. $8.75.

Lloyd-Jones, Esther, and Westervelt, Esther M. (eds.), *Behavioral Science and Guidance. Proposals and Perspectives.* New York, Teachers College, Columbia Univ., 1963. Pp. ix, 128. $2.50.

Luce, R. Duncan, Bush, Robert R., and Galanter, Eugene (eds.), *Readings in Mathematical Psychology.* New York, John Wiley & Sons, 1963. Pp. ix, 535. $8.95.

Luce, R. Duncan, Bush, Robert R., and Galanter, Eugene (eds.), *Handbook of Mathematical Psychology.* New York & London, John Wiley & Sons, 1963. Pp. xiii, 491. $10.50.

Marx, Melvin H., and Hillix, William A., *Systems and Theories in Psychology.* New York, McGraw-Hill Book Co., 1963. Pp. ix, 489. $8.95.

Mednick, Martha T., and Mednick, Sarnoff A. (eds.), *Research in Personality.* New York, Holt, Rinehart & Winston, Inc., 1963. Pp. x, 627. $6.50.

Miller, George A., *Language and Communication.* New York, Toronto & London, McGraw-Hill Book Co., 1963 (Copyright 1951), Pp. xiii, 298. $3.45. (paperback)

North, Robert C., Holsti, Ole R., Zaninovich, M. George, and Zinnes, Dina A., *Content Analysis.* Evanston, Ill., Northwestern Univ. Press, 1963, Pp. xiv, 182. $2.95. (paperback)

O'Connor, N., and Hermelin, B., *Speech and Thought in Severe Subnormality.* New York, Macmillan Co. (Pergamon Press), 1963. Pp. xi, 122. $5.50.

Oraison, Marc, *Illusion and Anxiety,* New York, Macmillan Co., 1963. Pp. xxix, 153. $3.95.

Orth, Charles D., *Social Structure and Learning Climate: The First Year at the Harvard Business Office.* Boston, Division of Research, Graduate School of Business Administration, Harvard Univ., 1963. Pp. xv, 236. $3.50.

Pavlov, Ivan P. (ed. and translated by W. Horsley Gantt), *Lectures on Condi-*

tioned Reflexes. New York, International Publishers, 1963. Pp. 414. $6.00. (Reissue, published 1928).

PEARSALL, MARION, *Medical Behavioral Science: A Selected Bibliography of Cultural Anthropology, Social Psychology and Sociology in Medicine.* Lexington, Ky., Univ. of Kentucky Press, 1963. Pp. viii, 134. $4.00.

QUAY, HERBERT S. (ed.), *Research in Psychopathology.* Princeton, N.J., D. VanNostrand Co., Inc., 1963. Pp. viii, 216. $1.95. (paperback)

ROBERTIELLO, RICHARD C., FRIEDMAN, DAVID B., and POLLENS, BERTRAM, *The Analyst's Role.* New York, Citadel Press, 1963. Pp. 126. $3.50.

ROLO, CHARLES (ed.), *Psychiatry in American Life.* Boston, Little, Brown & Co., 1963. Pp. viii, 246. $5.00.

RUCH, FLOYD L., *Psychology and Life.* (6th ed.), Chicago, Scott, Foresman & Co., 1963. Pp. 685. $8.00.

RUCH, FLOYD L., WARREN, NEIL, and GRAY, PHILIP H., *Working with Psychology.* (A Workbook for use with *Psychology and Life,* 6th ed.). Chicago, Scott, Foresman & Co., 1963. Pp. ix, 239. $2.25.

SHIPP, THOMAS J., *Helping the Alcoholic and His Family.* Englewood Cliffs, N.J., Prentice Hall, Inc., 1963. Pp. 140. $2.95.

SMELSER, NEIL J., and WILLIAM T. (eds.), *Personality and Social Systems.* New York & London, John Wiley & Sons, 1963. Pp. xi, 660. $8.95.

SOSKIN, WILLIAM F., *Verbal Interaction in a Young Married Couple. Social Science Studies.* Lawrence, Kan., Univ. of Kansas Publ., 1963. Pp. 25. (paperback)

SOUTHWICK, CHARLES H. (ed.). *Primate Social Behavior.* Princeton, N.J., D. VanNostrand Co., Inc., 1963. Pp. vi, 191. $1.95. (paperback)

STEPHENS, WILLIAM N., *The Family in Cross-Cultural Perspective.* New York, Holt, Rinehart & Winston, Inc., 1963. Pp. xv, 460. $4.75. (paperback)

STORR, ANTHONY, *The Integrity of the Personality.* Baltimore, Md., Penguin Books, 1963 (Copyright 1960). Pp. 186. $.95. (paperback)

TOMKINS, SILVAN S., and MESSICK, SAMUEL (eds.), *Computer Simulation of Personality: Frontier of Psychological Theory.* New York & London, John Wiley & Sons, 1963. Pp. x, 325. $5.00.

VERDIER, PAUL ANDRE, *Bio-psychology: A Biological Approach to Human Motivation.* New York, Exposition Press, 1963. Pp. 216. $6.00.

WATSON, GOODWIN (ed.), *No Room at the Bottom. Automation and the Reluctant Learner.* Washington, D.C., National Education Assoc., 1963. Pp. ix, 102. $2.00, paper, $3.00, cloth.

WEPMAN, JOSEPH M., and HEINE, RALPH W. (eds.), CONCEPTS OF PERSONALITY. Chicago, Aldine Publ. Co., 1963. Pp. xxix, 514. $8.95.

ZAX, MELVIN, and STRICKER, GEORGE, *Patterns of Psychopathology.* New York, Macmillan Co., 1963. Pp. 314. $3.50. (paperback)

INDEX

CHARLES VALLEY BROOK, University of Texas

AUTHORS

(The names of authors of original articles are printed in CAPITALS and SMALL CAPITALS; of authors of books reviewed, in roman; and of reviewers in *italics*.)

ALFERT, E. 638
ALLUISI, E. A. 512
Arnold, M. B. 516
Ausubel, D. P. 165

Baer, D. M. 173
BAKAN, P. 39
BATTIG, W. F. 66
BAXTER, R. 665
BEARE, A. C. 248
BERG, H. W. 492
BERRY, R. N. 333
BEVAN, W. 464
BIERSDORF, W. R. .. 1
Bijou, S. W. 173
BILODEAU, E. A. ... 293
BITTERMAN, M. E. 353, 366
Bjerstedt, A. 168
Blum, G. S. 526
BOARDMAN, W. K. . 503
BOLEN, C. H. 691
BORING, E. G. 336
Boring, E. G. 530
BOURNE, L. E., JR. . 229
BOWMAN, R. E. 436
BRACKETT, H. R. .. 66
Brazier, M. A. B.,... 524
BROOKS, V. 326
BROWN, C. C. 138
BRYAN, J. D. 82
BULLOCK, D. H. ... 353
Buros, O. K. 714
BURRIGHT, R. G. ... 205
Buss, A. H. 713

CARDNO, J. A. 509
CARP, F. M. 488
CAUTELA, J. R. 128
CHAMPION, R. A. 616, 679
CHENZOFF, A. P. .. 110
CLARK, B. 18
COLEMAN, E. B. ... 239
COPPOLA, L. A. 135
COSTELLO, C. G. ... .324, 670, 698
Cram, D. 712
CROSS, H. A. 318
CROVITZ, H. F. 631

DALLENBACH, K. M. 152, 158, 700
DAVENPORT, J. W. . 694
DAVIDSON, L. 191
DAWSON, W. W. .. 687
DENNER, B. 287
DICK, E. D. 607
DI LOLLO, V. 644
Dimmick, F. L., ... 519
DREGER, R. M. 703
Duijker, H. C. J. .. 172
DYAL, J. A. 332, 484
Dyk, R. B. 709

ELITHORN, A. 506
ELLIOTT, L. L. 589
Ellis, A. 169
ENGEN, T. 96
EPSTEIN, W. 257
ERICKSON, C. C. ... 458
ERVIN, S. M. 49
ESKILDSEN, P. R. ... 321
ESKIN, R. M. 366
EWART, A. G. 488

Faterson, H. F. 709
FILLENBAUM, S. ... 103
Forbes, T. W. 528
FRENCH, G. M. 446
Frijda, N. H. 172
Fuller, C. W. 521
FUREDY, J. J. 616
FURTH, H. G. 74

Gardner, R. W. 709
GIBSON, J. J. 386
GOGEL, W. C. 537
GONZALEZ, R. C. .. 366
Goodenough, D. R. . 709
GRAYBIEL, A. 18
GREGG, L. W. 110
Gruber, H. L. 159
Gulick, W. L. 531
GUNTER, R. C., JR. . 503
Hansel, C. E. M. 170, 170
HANSEN, B. 315, 492
HARKER, G. S. 537
Harris, D. B. 166
HARTMAN, A. M. .. 120
Harvey, O. J. 163
Haynes, R. 170
HEATHCOTE, M. J. . 679
HEINEMANN, E. G. . 653
HELD, R. 146
HELSON, H. 464
Henle, M. 163
HERRERA, B. 670
Heywood, R. 170
HINCKLEY, E. D. .. 116
HOCHBERG, J. 326
HOLLAND, H. C. ... 670
HOLLAND, T. A. ... 332
HOWARD, I. P. 150
HOWARD, T. C. 335
Hunt, D. E. 163

INGRAM, R. D. 458
IVERSON, M. A. 514

Jacobs, H. H. 528
Jensen, A. R. 166
JOHNSON, J. W. ... 293
JONES, D. 506

KAHNEMAN, D. ... 554
KANUNGO, R. 421
Karp, S. A. 709
KARSH, E. B. ...413, 429
KENNEDY, J. L. ... 704
KENYON, J. 140
KEPPEL, G. 470
KERR, M. 506
Kilpatrick, F. P. ... 159
KIMMEL, H. D. 665
Kenny, J. A. S. 519
Klaus, D. J. 712
Klausmeier, H. J. .. 165
Kluckhohn, F. R. ... 342
KOZIL, D. J. 1
KRAUSKOPF, J. 496
KRAVITZ, J. H. 554

LAMBERT, W. E. 421
LANDAR, H. 49
LAUGHERY, K. R. 110
LERNER, R. A. 683
Levine, E. S. 521
LINDAUER, J. 554
LINDNER, R. S. 687
LIPSCOMB, D. B. 631
LOGAN, F. A. 480
London, P. 526
LURIA, S. M. 598
Luria, S. M. 519
Lyle, J. 711

MacLeod, R. B. 532
MAIER, R. A. 464
MALMSTROM, E. J. 446
MARSHALL, A. J. 644
MARSHALL, M. A. 281
MARTIN, J. G. 675
MCFARLAND, J. H. 287
McGranahan, D. V. 172
MENDEZ, R. A. 74
MIKAELIAN, H. 128
MILAUCKAS, E. W. 82
MILLER, G. A. 157
MILLER, W. C. 515
MOORE, M. E. 191
MÜLLER, H. F. 140
MURRAY, C. S. 28

National Physics Laboratory 519
NEALEY, S. M. 329
NEISSER, U. 376
NELSON, J. T. 675
NEWMAN, E. B. 155
NICORA, B. D. 691
NOBLE, C. E. 579
NORRIS, C. M. 659
Nunnally, J. C. 714

OHWAKI, S. 1
Olson, W. C. 711
Orbach, J. 524
OVER, R. 452

PANGBORN, R. M. 315, 492
Parker, E. B. 711
PARKER, G. V. C. 579
Pauk, W. 171
PAUL, C. 124
PERLOE, S. J. 395
PERRY, N. 687
PICK, A. D. 386
PIERREL, R. 28

RETHLINGSHAFER, D. 116
RILEY, D. A. 329
Robertiello, R. C. 169
ROBERTS, W. A. 353
Robinson, F. P. 171
ROLLIN, A. R. 500
Rosenblatt, F. 705
ROSENBLUM, L. A. 318
RUDOLPH, R. L. 653
RUNQUIST, W. N. 281
Ryan, T. A. 161, 530

SALTZ, E. 124
Schramm, W. 711
Schroder, H. M. 163
SCHUCK, J. R. 135
SCHWARTZ, M. 266
Scott, W. A. 707
SHEPARD, R. N. 218
SHEPP, B. E. 500
Siegel, A. E. 173, 713
SLAMECKA, N. J. 274
STABLER, J. R. 484
Steiner, I. D. 168
STEPHEN, J. 324
STEVENSON, H. W. 299
STRAUGHAN, J. H. 148
Strodtbeck, F. L. 342
SUTHERLAND, N. S. 177
Symonds, P. M. 166

TANNENBAUM, P. H. 589
TEMPLETON, W. P. 150
TERWILLIGER, R. F. 306
THALLER, K. E. 500
THOMPSON, R. 39
THURLOW, W. R. 436
TROMATER, L. J. 205
TULVING, E. 143

UNDERWOOD, B. J. 470

Von Békésy, G. 531
Vygotsky, L. S. 532

WAGNER, A. R. 480
Walk, R. D. 707
WALLACH, H. 191, 404, 413, 429, 568, 659
Wallraff, C. A. 161
WAPNER, S. 287
WEIDENFELLER, E. W. 311
WEINSTEIN, S. 475
WEIR, M. W. 299
WEISS, J. H. 624
WERNER, H. 287
Werthelmer, M. 707
White, B. W. 705
WHITMAN, J. R. 138
WHITMAN, R. N. 124
WILSON, K. V. 89
WINNICK, W. A. 683
WIST, E. R. 537
Witkin, H. A. 709
WONG, R. 511

YOUNG, P. T. 205
Young, P. T. 516
YOUNG, R. K. 82, 458

ZIMMER, H. 691
ZIMNY, G. H. 58, 311
ZUCKERMAN, C. 404

SUBJECTS

(References in *italic* figures are to reviews.)

Abilities, human, learning and, *165-166.*
Accommodation, convergence, apparent size and, 1-17; distance and, 659-664.
Acuity, scotopic, body-position and, 598-606.
Adaptation, olfactory, 96-102.
Adaptation-level, judges' characteristics, influence of, on, 116-119; semantic restrictions vs., 103-109.
Adloescent, adult vs., modes of thinking, *166-168.*
After-effect, figural, tilt, a result of, 146-148; kinesthetic, interpolated stimulus, role of, in 670-674.
After-images, figural after-effects and, 306-310; interocular transfer, 308 f.; visual, apparent vs. target-size, 120-123.
Age, Gestalt laws and, 74-81; probability learning and, 299-305.
Aging, process of disengagement, *341-342.*
Aggression, psychology of, *713 f.*
Alpha conditioning (Planaria), 668 f.
Amherst College, psychological laboratory, 694-697.
Animals: Apparatus, Wisconsin general test for rats, 500-502
  Monkeys, neonatal, performance in visual cliff-situation, 318-320.
  Octupuses, shape-discrimination of, 177-190.
  Pigeon, resistance to extinction, 353-365; stimulus-generalization vs. discriminative training, 653-658.
  Planaria, conditioning and extinction in, 665-669.
  Rats, apparatus, consumption of liquids, electrically quantified, 687-690; dark-reared, 329-332; incentives, cognitive and S–R interpretation, 616-623; learning, S–R vs. relative-valence theory of, 679-682; resistance to extinction, 332 f.; taste-preferences, 205-217; visual occluders, contact, 135-137.
Aniseikonia, *160;* lenses, effect of, 539 f.
Anxiety, discriminative reaction-time and, 484-487.
Apparatus: Binary perceptual maze, 506-508.
  Brightness, scaling, 321-323.
  Centrifuge, human, 150-152.
  Color-mixer, monochromatic primaries, 496-499.
  Haploscope, 5 f.
  Liquid consumption (rats), electrically quantified, 687-690.
  Machine, conditioned reflex, 128-134.
  Maze, binary perceptual, 506-508.
  Measuring liminal movements, 138-139.
  Octupuses, shape-discrimination, 179.
  Psychological laboratory (Amherst College), 694-697.
  Sweat, finger, visual rating-scale of, 140-142.
  Tachistoscope, flexible four-field, 503-505; high-speed, constant-illumination, 691-693.
  Tape-recorder as a time-instrument, 324 f.
  Telestereoscope, 192.
  Visual, occluders, contact (rat), 135-137.
  Visual cliff, 318.
  Wisconsin general test for rats, 500-502.
Apparent size, after-images, 120-123; coins, monocular vs. binocular vision, 260; distance vs. assumed, 257-265; instructions, effect of, on, 1-17.
Association, backward vs. forward, 470-474; color-sweetness, 494; color-taste, 316 f.; constrained, informational uncertainty vs. rate, 218-228; statistical vs. Hull-Spence theory, 480-483; word-, Navaho, 49-59.
Astigmatism, scotopic, acuity and, 598-606.
Attention, kinesthetic after-effects and, 39-48.
Audition. *see* Music.
Autobiography, selected papers and, of, E. G. Boring, *530 f.*

Behavior, brain and, conference on, *524-526;* 'looking,' 386-394; stimulus-differentiation, 124-127.
Behavior, accident research, symposium on, *528-529.*
Binocular rivalry, dominance, nasal vs. temporal, 631-637; theories of, 631 f.
Biography, E. G. Boring, *530 f.; see* Necrology.
Blind, recognition of tactual forms, 488-491.
Blindness, occluders, 135-137.
Body-position, scotopic acuity and, 598-606.

Books received, *173-176, 348-352, 533-536, 716-718.*
Boring, E. G. Autobiography and selected papers, *530 f.*
Brain, behavior and, conference on, *524-526;* mechanisms, perceptrons and theory of, *705-707.*
Brightness, apparatus for scaling, 321-323.

Category-ratings, magnitude-estimations and, 395-403.
Centrifuge, human, 150-152.
Certainty, expectancy and, 333 f.
Character, national stereotypes, *172.*
Child, development, *173;* school-, world of the, *168-169;* television and the, *711 f.*
Choice, behavior, reinforcement in, 480-483; reaction-time, meaningful similarity and, 274-280.
Cognition, incentives and motivations of, *616-623;* letter sequences, 143-146.
Color, binocular rivalry, 631-637; sweetness-discrimination and, 492-495; symposium on, *519-521;* taste and, 315-317.
Color-mixer, monochromatic primaries, 496-499.
Color-name, luminance-level vs., 248-256; wave-length and, 248-256.
Concepts, formation, factors affecting, 229-238; generalization of, 607 f.
Conceptual systems, personal organization and, *163-165.*
Conditioned reflex, machine illustrating, 128-134.
Conditioning, animal vs. machine, 132; Planaria, 668 f.
Confidence, decision-time and, 28-38.
Consonants, English, confusion of, 89-95.
Constancy, depth-perception and, 404-412.
Content, informational vs. stimulus-material, 143-146.
Context, concept-formation and, 229-238; recall and, 239-247; of number and, 464-469.
Contours, figure-ground, 326-329.
Convergence, apparent size and, 1-17.
Cue, distance-, accommodation and, 659-664; depth-, 410.
Cutaneous sensitivity (pressure), dextral vs. sinistral *S*s, 475-479.

Deafness, *521-524;* rehabilitory techniques, *521-524.*
Decision-time, confidence and, 28-38; connotational meanings, of, 276-280; without reaction-time, 376-385.
Depth-perception: Anaglyphic, 408-410.
  Dark-reared rats, 329-332.
  Discrimination, neonatal monkeys, 318-320.
  Stereoscopic, constancy of, 404-412; modification of, 191-204, by the kinetic depth-effect, 429-435, of rapid, 413-420.
  Telestereoscopic, modification of, 413-420, 429-435, 436-445.
Determinants, positional effects in serial learning, 436-445.
Development, child, *173;* studies of, *709-711.*
Differentiation, psychological, studies of development, *709-711.*
Discrimination: Apparatus (rats) 500 ff.;
  Depth- (monkeys, neonatal), 318-320, (rats), 329-332; shape-, stationary (octupuses), 177-180;
  Reaction-time, anxiety vs. intelligence, 484-487;
  Sweetness, color, influence of, on, 494-495;
  Training, atimulus-generalization, effect of, on, 653-658;
  Two-point (breast), 475-479.
Distance, accommodation, cue of, 659-664; apparent, size and, 257-265; estimates, size vs., 452-457; formula for. 262, monocular vs. binocular, 260; perceived ratio to perceived size, 537-553.
Distraction, learning, effect of, on, 683-686, intentional vs. incidental, under, 683-686.
Dollars demanded for task-performance, questionary on, 293-298.
Drugs, stimulant and depressant, 698-700.
Duration, sensory-tonic theory, 287-292.

Educational psychology, learning and human abilities, *165-166.*
Emmert's law, after-images, 120-123; size, 306-310.
Emotion, personality and, *516-519.*
Emotional illness, treatment of, *169-170.*
English consonants, confusion of, 89-95.
Erratum, 158.
Estimates, linear extent of, effect of modulus on, 511 f.
Estimation, number (beans), 464-469, role of context in, 464-469.
Expectancy, certainty and, 333 f.
Experience, varied, function of, *339-341.*
Experimental psychology, 1797-1896, 509-511; results, methods of presenting, 700-702.
Extinction, resistance to, 332 f. (fish), 366-375 (pigeon), 353-365.

ESP, inquiries into, *170-171.*
Extravert-introvert, differences in persistency, 698-700.

Fantasies, adolescent vs. adult, *166-168.*
Figural after-effect, interocular transfer, 307 f.; after-images and, 306-310; characteristics of letters and numbers, 326-329.
Form, tactual, recognition of (blind vs. sighted), 488-491.
Freud, repression and defense, on, *337-339.*

Geriatrics, *341-342.*
Gestalt laws, 326-329; language and age, influence of, on, 74-81; principles, closure, figure, form, proximity, and similarity, 75-81.
Guessing, 'spread of effect' and, 638-643.

Hearing, experiments on, *531 f.*
Heart-rate, music and, 311-314.
Hedonic interactions, algebraic summation between sweet and bitter, 205-217.
Hering's illusion, illumination vs. distance, 644-217.
History, experimental psychology, 1797-1896, 509-511.
Horizontal, perception of, factors in the 18-27.
Hypothesis, size-distance invariance, 549-553.

Illness, emotional and psychiatric, *169-170.*
Illusion, Hering's impoverished, stimulus, 644-652; Oculogravic, 18-27.
Imagery, visual, tactual form-recognition and, 488-491.
Incentives, cognitive vs. S–R interpretation of, 616-623.
Industrial psychology, *333-334.*
Information-theory, letter-sequences and tachistoscopic identification, 143-146.
Informational uncertainty, rate of association and, 218-225.
Instruction, judgment of size and, 1-17.
Intelligence, discriminative reaction-time and, 484-487.
Introvert-extravert, persistency differences, 698-700.

Judgment: Age, of influence of judges' age on, 116-119.
  Categorical vs. ratio, 103-109.
  Category-ratings vs. magnitude-estimations, 395-403.
  Color-names at two luminance-levels, 248-256.
  Confidence, decision-time and, 28-38.
  Contextual effects from language-restrictions, 103-109.
  Height, of, influence of judges' height on, 116-119.
  Horizontal and vertical, 18-27.
  Instructions, effect of, on, 15 f.
  Semantic differential of nonsense-shapes, 589-627.
  Similarity, olfactory, 99 ff.
  Size, apparent, 1-17; distance, 543-546.
  Under reduced and unrestricted viewing conditions, 452-457.
  Weights, lifted, 28-38; restricted viewing conditions, 452-457; weights, lifted, 28-38.

Kinesthetic after-effects, attention and, 39-48; interpolated stimulus, role of, in, 670-674; satiational vs. frame of reference theory, 670-674.
Kinetic depth-effect, 191-204; stereoscopic vs., 429-435.

Laboratory, multipurpose (Amherst), 649-697.
Language, influence on Gestalt laws, 74-81; phonetic symbolism in, 624-630; thought and, *532.*
Leadership, organization and, *347.*
Learning: All-or-none, 675-678.
  Conceptual, similarity and training, function of, 607-615.
  Differentiation vs. generalization, 124-127.
  Discriminative, apparatus for (rats), 500-502; training vs. stimulus-generalization, 653-658.
  Distraction, under, 683-686.
  Human abilities and, *165-166.*
  Intentional vs. incidental, 683-686.
  Machine, conditioning, value of, 133 f.
  Method, serial, 82-88; single-trial vs. incremental, 675-678; spelling vs. pronouncing, 110-115.
  Paired-associate, 66, 73, 82-88, 281-286; bi- vs. unidirectional, 470-474; meaningfulness in, 579-588; performance in the runway, 124-127; presentation-rate, effect of, on, 458-463; prior serial learning, effect of, on, 458-463; rate of presentation, substitution, and mode of response on, 110-115.
  Probability-, age and verbalization in, 299-305.
  Relative-valence theory of, 679-682.
  Repetition, role of, in, 148 f.
  S-R vs. cognitive interpretation, 679-682.

  Serial, positional effects in, 436-455.
  Single-trial vs. accumulative, 110-115; incremental, 66-73, 266-273; repetition, 148 f.; serial, 82-88.
  'Spread of effect,' 638-643.
  Stimulus-generalization and, 607-615.
Limen, two-point (breast) dextral and sinistral *Ss*, 475-479.
Linear extent, estimates of, effect of modulus on, 511 f.
Liquid consumption (rats), electrically quantified, 687-690.
Looking, act of, 386-394; eye- and head-positions in, 387-394; verbal variants of 386.
Lotsof's certainty and expectancy, relation between, 233 f.

Magnitude, estimates of, category-ratings and, 395-403; effect of modulus on, 511 f.
Management, organization and, *347*.
Manifest anxiety, discriminative reaction-time and, 484-487.
Mathematical probability, psychological vs., 335.
Maturation, perception vs., 74-81.
Maze, a binary perceptual, 506-508.
Meaningfulness, paired-associate learning and, 579-588; semantic satiation and, 421-428.
Meetings: Amer. Philos. Soc., 336.
  Amer. Psychol. Assn., 155.
  Eastern Psychol. Assn., 514 f.
  Nat. Acad. Sci., 336.
  Rocky Mt. Psychol. Assn., 515.
  Soc. Exper. Psychol., 704.
  Sou. Soc. Philos. Psychol., 512-514.
  Southwestern Psychol. Assn., 703.
Memory-span, words, approximations to English vs., 239-247.
Mental hygiene, *344-345*.
Method: Cross-adaptation, smell, 96-102.
  Effective study, of, *171-172*.
  Learning, paired-associate, backward vs. forward, 470-474; meaningfulness in, 579-588; single-trial vs. incremental, 266-273, 675-678; *see* Learning: Paired-associate, serial, single-trial.
  Paired-associate learning, anticipatory interval, synonymity, and transfer, in, 281-286.
  Psychophysics, single stimuli, 103-109, 116-119.
  Questionary on tasks, task-intensity, and dollars demanded for performance, 293-298.
  Thematic appreception, *345-346*.
Mind, model of, a, *526-528*.
Mirror-drawing, direction vs. speed, 58-73; practice-effects in, 61-64.
Model, mind, of the, *526-528*.
Monkey, *see* Animals: Monkeys.
Motivation, cognitive vs. S–R interpretation of, 616-623.
Movement, apparatus for measuring, 138-139.
Music, galvanic skin-response and heart-rate, 311-314.

National character, stereotypes, *172*.
Navaho, word-association, 49-57.
Necrology, Fisher, R. A., 157-158.
Neurodynamics, principles of, *705-707*.

Obituary, *see* Necrology.
Occupational prestige, category-ratings vs. magnitude-estimation of, 395-403.
Octupuses, shape-discrimination of, 117-190.
Oculogravic illusion, contributing factors, 18-27.
Oculomotor adjustments, apparent size and, 1-17.
Olfaction, *see* Smell.

Pain, scale of, 152-155.
Paired-associate learning, *see* Learning: Paired-Associate; Method: Learning.
Perception: Age, role of, in, 74-81.
  Apparent size, 1-17.
  Audition, hearing, experiments on, *531 f.*
  Children's, of selves and of others, *168-169*.
  Deaf persons, of, 74-81.
  Depth-, dark-reared rats, 329-332; sterescopic, 191-204, 404-412; *see* Depth-Perception.
  Distance, ratio of, to size, 537-553.
  Gestalt laws of, 74-81.
  Horizontal and vertical, 18-27.
  Language, role of, in, 74-81.
  Looking behavior, 386-394.
  Psychology of, epistemology and the, *161-163*.
  Size-, ratio of, to distance, 537-553; and distance, relation between, 257-265.
  Space, *159-161*.
  Tilt, 146-148; centrifugal vs. gravitational effects, on, 150-152.
  Time, rhythmic activity and, 287-292.
  Visual, apparent size of after-images, 120-123.
  See Illusions.
Perceptrons, *705 f.*
Persistency, differences in, extravert-in-

trovert and sex, 698-700; physical, drugs and, 698-700.
Personality, conceptual systems and, *163-165;* emotion and, *516-519.*
Phenomenal object, 326-329.
Philosophical theory, psychological fact, and, *161-163.*
Phonetic symbolism, semantic differentiation and, 624-630.
Photopic vision, illusion (Hering's) and, 644-652.
Pigeon, *see* Animals: Pigeon.
Planaria, conditioning and extinction in, 665-669.
Practice-effects, mirror-drawing, 61-64.
Pressure-sensitivity (breast), dextral vs. sinistral *S*s, 475-479.
Probability, learning, age and verbalization, role of, in, 299-305; mathematical vs. psychological, 335.
Problem-solving, reinforcement in, 480-483.
Psychological laboratory (Amherst College), 649-697.
Psychology, aggression of, *713 f.*; experimental, 1797-1896, 509-511.
Psychophysical, judgments in weight-lifting, 28-38.
Pupil-size, apparent size vs., 1-17.
Puzzle-figures, 326-329.

Questionary, dollars demanded for task-performance, 293-298.

Rating-scale, sweat, finger, visual, 140-142.
Rats, *see* Animals: Rats.
Reaction-time, choice, meaningful similarity and, 274-280; decision-time without, 376-385; uncertainty, degree of, vs., 277-280.
Recall, English, approximations to, vs. words recalled, 239-247.
Recognition, tactual forms, sighted and blind *S*s, 488-491; test, single-trial vs. incremental learning, 675-678.
Reflex, condition, machine, 128-134.
Reinforcement, extinction and, (fish), 366-375; (pigeon), 353-365.
Repression, defense and, Freud's concept, *337-339.*
Research, accidents, on, symposium, *528-529;* psychological, introduction to *707-709.*
Response-latencies, differential semantic ratings, of, 277-280.
Retinal rivalry, binocular, 631-637.
Rhythmic activity, time, perception of, and, 287-292.
Rivalry, retinal, binocular, 631-637.

Satiation, attention and, 39-48; semantic, meaningfulness and, 421-428; theory, kinesthetic after-effects and, 670-674.
Scales, category-ratings and magnitude-estimations, 395-403; pain, 152-155; symmetry, semantic differential and, 446-451.
Scaling, age, height and, 110-119; analysis, multidimensional, 89-95; factor-matrix, olfactory similarity-judgments, 101.
Scanning-time, 376-385; context and, 381 ff.; critical item, position of, and, 377; practice and, 383-385; spacing, horizontal, and 381.
Scotopic vision, illusion (Hering's), under, 644-652.
Self-perceptions, children's, *168-169.*
Semantic differential, connotative meaning and the, 554-562; phonetic symbolism and, 624-630; scale-symmetry and, 446-451; structure of inferences and the, 554-567.
Semantic differential responses, nonsense-shapes, to, 589-627; prediction of factor-scores from stimulus-shapes, and, 583-627.
Serial learning, effects of, in, position, 436-445; list vs. trial-position in, 436-445.
Set, judgment of size, effect of, on, 6 ff.; kinesthetic after-effects and, 670-674.
Shape, discrimination, octopuses, of, 177-190.
Similarity, concept-formation, function of, 607-615; response-latencies, 274-280.
Single-stimuli, *see* Method: Psychophysics.
Single-trial learning, recognition-tests and, 675-678; *see* Learning: Single-trial.
Size, after-images, real and apparent, 120-123; apparent, *see* Apparent size; distance and, relation between, 257-265; perceived, method of measuring, 542, ratio of, to perceived distance, 537-553.
Size-constancy, 538 f.
Size-distance invariance, 537-543.
Size- and distance-estimates, reduced and unrestricted viewing conditions, under, 452-457.
Size-perception, determinant of distance-cue, 660 f.
Smell, adaptation, aliphatic alchools, 96-102.
Social behavior, *347-348.*
Social psychology, cross-cultural gener-

alizations, *342-343*; introduction to, *346-347*.
Space, perception, *159-161*.
'Spread of effect,' guessing and the, 638-643.
Stereoscopic depth, constancy of, 404-412; kinetic depth-effect and, 429-435; modification of, rapid, 413-420; perception, 191-204; 318-320; 329-332.
Stimulus-differentiation, runway, in, 124-127.
Stimulus-generalization, discriminative training and, 635-658.
Stimulus-response interpretation, incentives and motivations of, 616-623.
Stimulus-scaling, multidimensional, 89-95.
Study, effective, *171 f*.
Sweat, finger, visual rating scale of, 140-142.

Tabular vs. graphic methods of presenting results, 700-702.
Tachistoscope, flexible four-field, 503-505; high-speed, constant-illumination, 691-693.
Tactual forms, recognition of, by blind and sighted *Ss*, 488-491.
Tactual recognition of form, blind vs. sighted *Ss*, 488-491.
Task-performance, dollars demanded for, 293-298.
Taste, rat's preferences for sucrose and quinine hydrochloride, 205-217; sour and sweetness, color, effect of, on, 315-317; sweetness, discrimination, dry wine, influence of color on, 492-495.
Teaching machines, explaining and programming, *712 f*.
Telestereoscope, 413-420, 429-435.
Television, impact on, children, *711 f*.
Tests in print, list and short description, *714 f*.
Thematic apperception methods, issues in, *345-346*.
Thematic apperceptions, adolescent vs. adult, *166-168*.
Thinking, adolescent vs. adult, *166-168*.
Thought, language and, *532*.
Threshold, taste, aversion, quinine vs. sucrose (rat), 208 f., quinine vs. water (rat), 207 f.; visual acuity, body- and head-position vs., 598-606.
Tic douloureux, pain of, 152-155.
Time, perception of, organismic theory, 287-292.
Time-instrument, tape-recorder as a, 324 f.
Training, discriminative, stimulus-generalization and, 653-658; transfer of, 82-88.
Transactional psychology, *159-161*.
Transfer effects, anticipatory interval, synonymity in paired-associate learning, 281-286; serial to paired-associate learning, 458-463.

Uncertainty, informational, rate of association and, 218-225.

Value-orientation, variations in, *342-343*.
Verbalization, probability-learning and, 299-305.
Vertical, perception of, factors in the, 18-27.
Vision: Adaptation to displaced visual direction, 568-578.
  After-image, figural after-effects and, 306-310.
  Binocular vs. monocular, 7 f.
  Brightness, apparatus for scaling, 321-323.
  Color, problems of, symposium on, *519-521*.
  Contact occluders (rats), 135-137.
  Discrimination, dark-reared rats, 329-332.
  Illusion, oculogravic 18-27; scotopic, vision (Hering's), under, 644-652.
  Line of regard, 386-394.
  Oculomotor adjustments and apparent size, 1-17.
  Photopic, illusion (Hering's), under, 644-652.
  Retinal rivalry, bionocular, 631-637.
  Scotopic, acuity, body-position, effect of, on, 598-606; illusion (Hering's), under, 644-652.
  Wave-length vs. color-name, 248-256.
Visual, after-images, Emmert's law and, 120-123.
'Visual cliff," depth-discrimination (rat), 329-332.

Wave-length, color-name and, 248-256.
Weight-lifting, judgment in, 28-38.
Wisconsin general test for discrimination (rats), 500-502.
Word-association, commonality, Navaho, 49-59; paradigmatic responses, 49-59.
Words, sound and meaning of, 624-630.

GENL. & PSYCL. LIBRARY